International Journal of Robotics and Control Systems
Vol. 4, No. 3, 2024, pp. 979-999
ISSN 2775-2658
http://pubs2.ascee.org/index.php/ijrcs

Comparative Study of ANN and SVM Model Network Performance for Predicting Brake Power in SI Engines Using E15 Fuel

Mohamed S. Hofny [a,1], Nouby M. Ghazaly [a,b,2,*], Ahmed N. Shmroukh [a,c,3], Mostafa Abouelsoud [a,4]

[a] Faculty of Engineering, South Valley University, Qena-83523, Egypt
[b] Technical College, Imam Ja'afar Al-Sadiq University, Baghdad, Iraq
[c] Faculty of Industry and Energy Technology, New Cairo Technological University, Cairo 11835, Egypt
[1] Mohamed.Salah@eng.svu.edu.eg; [2] Nouby.Ghazaly@eng.svu.edu.eg; [3] eng_ahmednagah@yahoo.com;
[4] mostafa.soud@eng.svu.edu.eg
* Corresponding Author

ARTICLE INFO

Article history
Received May 02, 2024
Revised June 11, 2024
Accepted June 16, 2024

Keywords
Artificial Neural Network;
Engine Performance;
SI Engines;
Support Vector Machine;
Machine Learning

ABSTRACT

Currently, artificial neural networks (ANNs) and support vector machines (SVMs) are the most common applications of machine learning approaches. In this study, a comparative study of ANN and SVM is presented to evaluate the performance of each model in predicting the brake power (BP) of GX35-OHC 4-stroke, air-cooled, single cylinder gasoline engine with E15 (15% ethanol + 85% gasoline) fuel. Two models are compared based on experimental dataset that has single output (BP) and five inputs, engine speed (S), engine torque (T), intake air temperature (T_{air}), intake air flow (Q_{air}), and fuel consumption ($\dot{m}$). The samples were split into three sets: Training set (70%), Validation set (15%), and the Test set (15%) based on 60 samples. The results are compared through different graphs such as target vs actual values, regression plots, histograms of prediction errors, residual plots, learning curves, and error distributions. The results showed that SVM model is indicated to have lower RMSE (0.0044) and higher EVS (0.9953), while ANN is shown to have lower value of MAPE (1.51%). These results have significant implications for the use of ANN and SVM models in real-world applications that need gradual comprehensibility and model generalization. In addition, work done with the models outlined above should try and test them in other engines and operating conditions to demonstrate the model's and performance.

1. Introduction

Artificial intelligence (AI) has significantly progressed over the past decade, as it is one of the most common techniques used in solving complicated problems [1]-[4]. Machine learning (ML) is one of the most prevalent branches of AI and has extensive applications across numerous fields [5], [6]; it learns from examples or data provided and then is used for predicting new cases [7]. The use of ML algorithms includes many areas, such as automation, robot recognition patterns, prediction, healthcare, energy, and manufacturing [8], [9], and the prediction of the performance and emissions of IC engines [10]. There are various ML algorithms used for prediction purposes, such as SVMs, ANNs, kernels and nearest neighbors (k-NNs), and deep learning (DL) [11], [12]. In the field of IC engines, ANN and SVM algorithms are commonly used for predicting engine performance and emissions [13]-[15].

Artificial neural networks are among the most common techniques used in the prediction of ICE performance [16]-[18]. They are considered the core of DL algorithms. They consist of three layers, an input layer, one or more hidden layers, and an output layer [19], [20]. They draw inspiration from the workings of biological neurons in the nervous system, which consists of numerous interconnected neurons that function together. This makes ANNs useful for solving complex, highly nonlinear, and massively parallel problems. The layers consist of neurons that are connected from one layer to another through various weights. This architecture, which features many neurons arranged in different layers, can be trained or configured to carry out specific tasks by carefully adjusting its structure, biases, and connection weights [21], [22]. These models have been used to predict engine parameters such as torque, combustion chamber pressure, exhaust gas physical properties, and vibration. Additionally, ANNs have been employed for fault detection, misfire diagnosis, and combustion timing control in IC engines [23]-[26]. ANNs are more effective than chemical kinetic models and CFD because they can provide accurate predictions in less time and with fewer resource requirements [24], [27]. Furthermore, ANNs are used to predict the performance, emission, and combustion characteristics of IC engines operating on alternative fuels such as biodiesel and ethanol-diesel blends [28], [29]. The effectiveness of ANN models in predicting engine parameters has been demonstrated, with RMSEs in the range of 0.4-1.8% for engine performance prediction [30]. The support vector machine (SVM) is considered a strong mathematical tool used for classification, regression, and function estimation. It has also been applied in modelling machining operations. In SVMs, different types of kernel functions play an important role in training parameters [31], [32]. These include linear, polynomial, radial basis function (RBF), sigmoid, and Gaussian kernel functions [31], [33]. This technique is a supervised, nonparametric method in statistical learning that is known for its strong balance between accurately predicting outcomes and generalizing these models to new, unseen data. The strengths of SVM models lie in their ability to manage spaces with many variables and to effectively handle patterns with noise and varied distributions of properties such as those found in soil [34]. The SVM model is a powerful algorithm for performance and emissions prediction because it provides high accuracy [35]. Additionally, the single-Wiebe function has been shown to be useful in engine research practice, particularly when a simple combustion model with high computational efficiency is required as an input into more complex models [36]. SVMs have been used in the prediction of NOx emissions and brake mean effective pressure in diesel engines and have achieved high-performance results [37], [38]. SVMs have also been applied in the prediction of ICE output power. This demonstrated that SVMs exhibit considerable performance in predicting engine performance and reducing emissions [39], [40].

There is a comparative study of SVM and ANN in IC engines in the evaluation of machine learning performance in predicting various parameters. SVMs can drive decision functions and have been applied in diverse fields, such as challenges associated with ICE performance optimization, control and fault diagnosis in small piston engines [41]-[43]. Both SVMs and ANNs have been used in the ICE approach for various prediction purposes, such as engine performance and emissions [44]-[46]. The use of SVM and ANN algorithms in IC engines for various purposes, such as performance prediction, fault detection, and emission analysis, is very common. For example, an ANN is used in a diesel engine using waste cooking biodiesel fuel to successfully predict BP, torque, SFC, and exhaust emissions [47]. The choice between SVMs and ANNs in the application of IC engines depends on the specific task and the nature of the data. For example, the predictive power of SVM algorithms was compared with that of traditional ANN models, and it was shown that for a limited amount of experimental data, SVM has better performance in finding the global optimum solution than does ANN [38], [48]. Chowdhury [49], compared the accuracy and reliability of various ML and DL models, including random forest (RF), SVM, and ANN. The analysis of the performance of different algorithms, namely, ANN, RF, and SVM, showed that ANN had the highest overall accuracy and kappa coefficient compared with the other two algorithms. According to the study, the SVM models are ideal for handling heterogeneous and homogeneous land feature type data, while in noisy data where the target classes tend to overlap, the model has low performance. On the other hand, ANN models yielded higher accuracy in the classification of urban LULC because of the capacity to analyse noise and generalize relationships. Also, Kurani et al. [50], showed that ANN and SVM models are more suitable for application in stock prediction. ANN models, due to their ability to handle missing data points, have enhanced prediction efficiency through other ANN models, such as ANN-MLP and GARCH-MLP, along with backpropagation and multilayer feed-forward efficiency. Algorithms such as the simple boundary SVM have also provided better accuracy rates between 60% and 70% and can be combined with other algorithms such as the random forest algorithm and genetic algorithms. Almansour et al. [51], used ANNs and SVMs to classify patients into diseased and disease-free groups, and the data used consisted of 400 patients with 24 diagnostic factors. The missing attribute values were estimated by the average of the respective attributes; the optimal parameters of the ANN and SVM were set up by using cross-validation and several trials. An analysis of the findings of this study revealed that the ANN achieved an accuracy of 99%. 75% show that the proposed approach outperforms the other classification algorithms, including SVM, which had an accuracy of 97.75%. Therefore, more research concerning ANNs is needed for early CKD prediction. Li et al. [52], explored the application of support vector machine (SVM) and artificial neural network (ANN) methods and compared their effectiveness in assessing the vulnerability of urban buried gas pipeline networks. SVM demonstrated better performance, with a mean squared error (MSE) of 2.74E-4, compared to ANN's MSE of 1.92E-2. The SVM model achieved a symmetric mean absolute percentage error (SMAPE) of 0.79%, outperforming the ANN model, which had a SMAPE of 8.64%. Sharma et al. [53], predicted the mutagenicity of compounds and prevented costly drug failures during late development or clinical trials using SVM, ANN, and Bayesian classifiers. The classifiers were trained and tested on a dataset comprising compounds with known mutagenic properties utilizing seventeen descriptors. The SVM classifier with an RBF kernel achieved the highest overall prediction accuracy of 71.73% when the ANN-based classifier showed a lower accuracy of 59.72%, while the Bayesian classifier achieved an accuracy of 66.61%.

According to previous studies, both ANNs and SVMs are used in IC engines for the prediction of performance and emissions. However, it is necessary to investigate the application of both SVMs and ANNs in the context of IC engines. In this study, both the ANN and SVM models for predicting the BP of a GX35-OHC 4-stroke, air-cooled, single-cylinder gasoline engine using E15 fuel is conducted. The assumptions are that both models, SVM and ANN, will have the ability to estimate the engine performance; however, SVM is accurate if the sample size is significantly smaller, and ANN might perform better if the relationships in the data are nonlinear. In additions, this study analyses the performance of the forecasting model, ANN, and SVM based on parameters such as the RMSE, EVS, and MAPE; the advantages and disadvantages of each model are specified by performing detailed comparisons, such as comparisons of target vs. actual values, regression plots, histograms with prediction errors, residual plots, learning curves, and error distributions; and the limitations of these models and the ways in which they can be applied in practice as well as the need to establish model interpretability and extrapolation are discussed.

2. Method

2.1. Experimental Setup

The experiment is performed on a GX35-OHC 4-stroke, air-cooled, single-cylinder gasoline engine with the specifications shown in Table 1 and shown in Fig. 1, which describes all the components used in the study. The engine is matched with an HM-365 dynamometer to measure BP and fixed on a CT-159 unit equipped with various measurement devices, including temperature and flow sensors. The CT-159 unit includes an air tank equipped with an intake air temperature sensor and an air flowmeter. The input parameters for the two models include engine speed (measured by an optical sensor), engine torque (measured by a force sensor), intake air temperature (measured by the intake air temperature sensor), intake air flow (measured by the air flowmeter), and fuel consumption (measured by an electronic pressure sensor).

2.1.1. Experimental Procedure and Measurements

The procedures of the experiment for both the ANN and SVM (shown in Fig. 2) began with passing the air through an air blower that pushes fresh air into the connecting pipe where air flows through a flow control valve for airflow adjustment and passes through the air heater to increase the inlet air temperature. The air then passes through a filter and a flowmeter, which measures air consumption in litres per minute. This air is then directed through the carburetor, where it is mixed with a measured amount of fuel (E15) before entering the combustion chamber. Concurrently, there is a measurable decrease in the fuel level within the fuel measurement tube. This fuel consumption is either observed directly from the tube or calculated via a connected PC. The engine's operational parameters, such as rotational speed and brake torque, are monitored by an electronic dynamometer linked to the engine via a v-belt. This setup helps in calculating the required BP for the engine. The experimental results, which include BP (derived from engine speed and brake torque), are displayed on the PC due to the five input parameters.

$$\bar{X} = \frac{\sum Xm}{n} \tag{1}$$

Table 1. Engine specifications used in experiment

Engine Type	4-stroke single cylinder air cooled OHC petrol engine
Bore X Stroke (mm)	39 x 30 mm
Compression ratio	8.0: 1
Ignition System	Transistorized
Net Power	1.0 kW (1.3 HP) / 7000 rpm
Oil Capacity	0.1 Liter
Starting System	Recoil
Displacement	35.8 cm3
Fuel cons. at cont. rated power	0.71 L/h - 7000 rpm
Max. net torque	1.6 Nm/ 5500 rpm
Idle speed	2800RPM
Lubrication	Oil mist
Carburetor	Ruixing Brand Carburetor

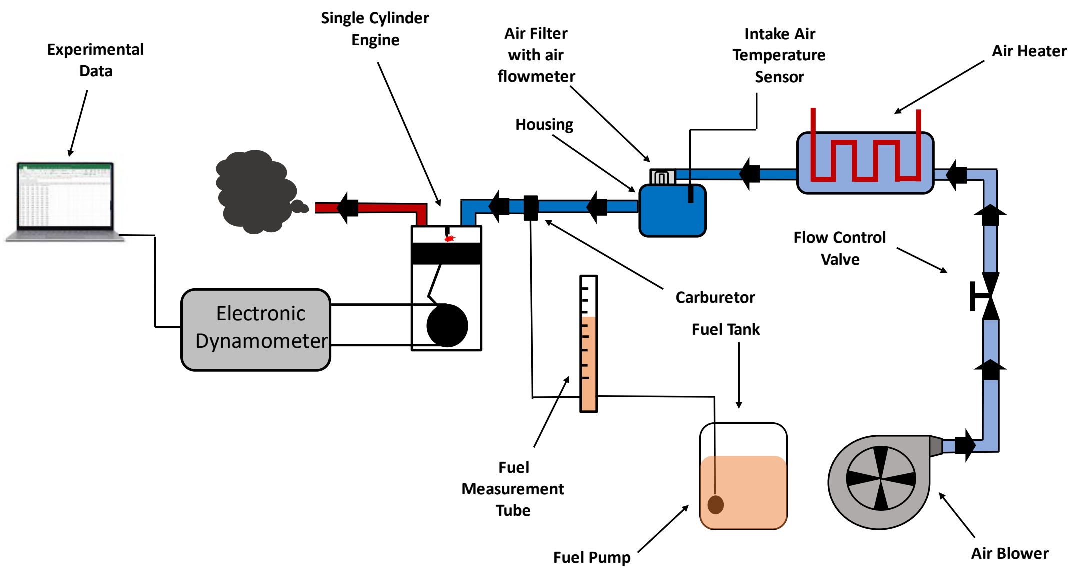

Fig. 1. Experimental installation cycle for ANN and SVM results

2.1.2. General Settings for Both ANN and SVM Algorithms

For instance, measurement errors associated with load, speed and temperature can introduce bias in the desired results. For example, changes in the load that is applied through the dynamometer or changes in the speed of the engine will interfere with the precise calculation of BP. Multiple measurements are conducted to determine the "n" number of experimental parameters. This collection of measurements is used to estimate the essential experimental outcome. Thus, eq (1) can be used to determine the mean of the amount measured in the studies [54]. Xm is the measured value, and n is the number of measurements. The formula for the standard deviation (SD) is given in Equ. (2), and the uncertainty (U) can be calculated by Equ. (3) As indicated in Table 2 [55], [56]. These uncertainties may lead to variations in the experimental results and the predictions of the ANN and SVM models. For instance, if the size and speed of the load differ significantly, it results in deviations in BP calculations and, therefore, in the model's performance measurement.

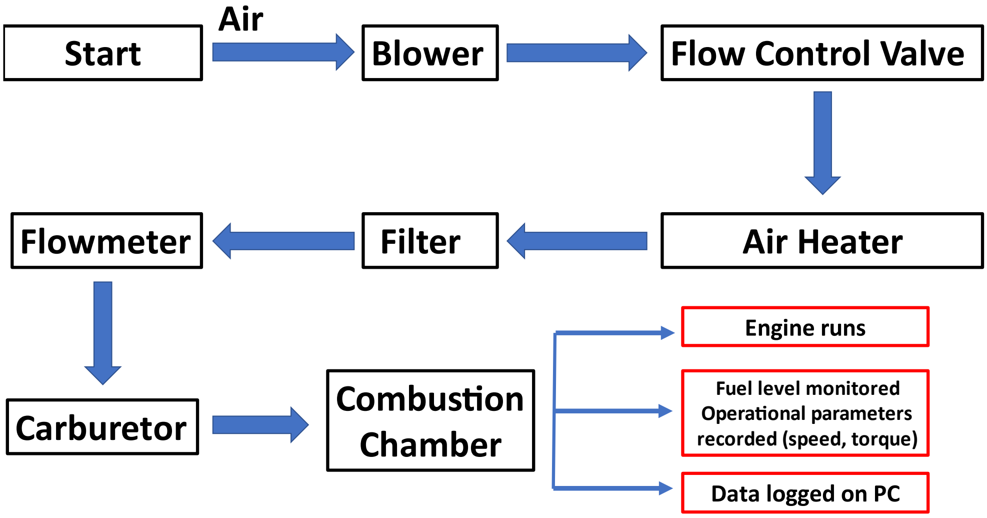

Fig. 2. Flowchart of experimental procedure

Table 2. Uncertainty measurements

Measurement	Uncertainties
Load	± 0.01 N
Speed	± 10 rpm
Temperature	± 1°C
BP	±1.60

$$SD = \sqrt{\frac{\sum_{m=1}^{n}(Xm - \bar{X})^2}{(n-1)}} \tag{2}$$

$$U = \frac{SD}{\sqrt{n}} \tag{3}$$

For applying the ANN and SVM models, the input data included the engine speed, engine torque, air temperature, air flow, and fuel consumption. The output parameter is BP. The samples used in both models are 60 samples with 70% training, 15% validation and 15% validation; this distribution is commonly used in various previous studies [50], [57], [58]. Methodological concerns that explain why 60 samples are adequate for capturing the sampling distribution of the statistic and handling the noise that tends to affect the training of models. First, the selection of features was based on prior knowledge of the domain, and the characteristics that most affect the engine parameters were selected. These were the engine speed, engine torque, intake air temperature, intake air flow and fuel consumption rates, which were tabulated and compared. Correlation analysis was performed to identify and remove highly correlated features, preventing multicollinearity. The samples were normalized to a range of 0.1 to 0.9 using min–max scaling to ensure that all features contributed equally to the model's decision-making process. To indicate the performance of the ANN and SVM models, we employed several metrics: RMSE, EVS, and MAPE. All of these factors define different aspects of the model and help to view the results from different perspectives.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{4}$$

$$EVS = 1 - \frac{Var(y_i - \hat{y}_i)}{Var(y)} \tag{5}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{6}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, n is the number of observations, $Var(y_i - \hat{y}_i)$ is the variance of the prediction errors, and $Var(y)$ is the variance of the actual values.

In general, the RMSE is the overall average of the errors that are committed with reference to the actual values and the predicted values. A lower RMSE value is better when it is in the context of errors; in other words, smaller errors denote a better model. EVS represents the percentage strength of the relationship between the dependent variable and the model being tested. It varies from 0 to 1, where any value close to 1 indicates that there is a high level of fit whereby the model contains most of the variability in the data. The MAPE quantifies predictions in percentage terms by looking at the absolute error made. It shows the average percentage deviation between the expected and real values. The MAPE is commonly used as the mean absolute percentage error, and its lower value is desirable because it allows easier judgement of the quality of the model in relative terms.

2.2. ANN Model Setup

The network's input layer receives several parameters that influence engine performance. These parameters are the engine speed (S), engine torque (T), intake air temperature (T_{air}), intake air flow (Q_{air}), and fuel consumption ($\dot{m}$). The network's sole output parameter is brake power (BP), as illustrated in Fig. 3. In the hidden and output layers, each neuron's output is determined by two functions: summation and activation. The summation function computes a weighted total of the inputs from the preceding layer. This weighted sum is then transformed by an activation function to produce the neuron's output. The summation function is used to sum the input values according to their weights with the addition of a bias term, as shown in eq (7). The sum is then used for the activation function input. The activation function plays an important role in ANNs because it adds nonlinear characteristics to the neuron's output by computing a neuron's output by processing the weighted sum of its inputs. The selection of a suitable activation function is important and relies on the specific problem being addressed and the expected range of outputs. The logistic sigmoid function is particularly popular in multilayer perceptron models because of its differentiable, continuous, and nonlinear nature [59], [60]. The sigmoid activation function is shown in eq (8). Normalization is performed by limiting both the input and output data to fit within a specific range, often between 0.1 and 0.9. The sigmoid activation function is adjusted using a designated formula, known as Eq. 6, in this context [61]. For regression tasks, the feedforward backpropagation network is a widely used form of ANN [62]. When applying an ANN, the data flow one way, starting at the input layer and ending at the output layer. The backpropagation algorithm is applied during the training process to fine-tune the neuron weights and biases to reduce the difference between the predicted and true outcomes [63]. The Levenberg–Marquardt (LM) backpropagation algorithm is used to reduce the error by aligning the predicted outputs with the actual outputs. The RMSE, R-squared, and MAE are used to indicate the performance of the ANN. The needed target is to reach the lowest RMSE and MAE and the highest R-squared values across BP predictions. The network's hidden layer is initially set with 3 to 12 neurons to find the ideal number. The hyperparameters that are commonly chosen for the ANN model. For BP prediction, 3 neurons in the hidden layer yield the best results in terms of accuracy, variance explained, and the absolute difference between the predictions and actual values, as shown in Table 3. A sigmoid function is employed as the activation function because of its ability to calibrate nonlinearity and because of its frequent use in multilayer perceptron models. The learning rate is fixed at 0.001, [64] to ensure stable learning and convergence, with a momentum constant of 0.1 to accelerate convergence and avoid local minima. The number of epochs is adjusted for each neural network to ensure sufficient training without overfitting, meaning that the network's weights and biases are updated during training based on these settings. An error goal of 1e-30 is set, along with a regularization parameter of 0.01, to avoid overfitting by penalizing large weights.

$$y = \sum_{i=1}^{n} (w_i \times x_i) + b \tag{7}$$

$$y = \frac{1}{1 + e^{-x}} \tag{8}$$

$$x_i = \frac{0.8}{d_{max} - d_{min}} (d_i - d_{min}) + 0.1 \tag{9}$$

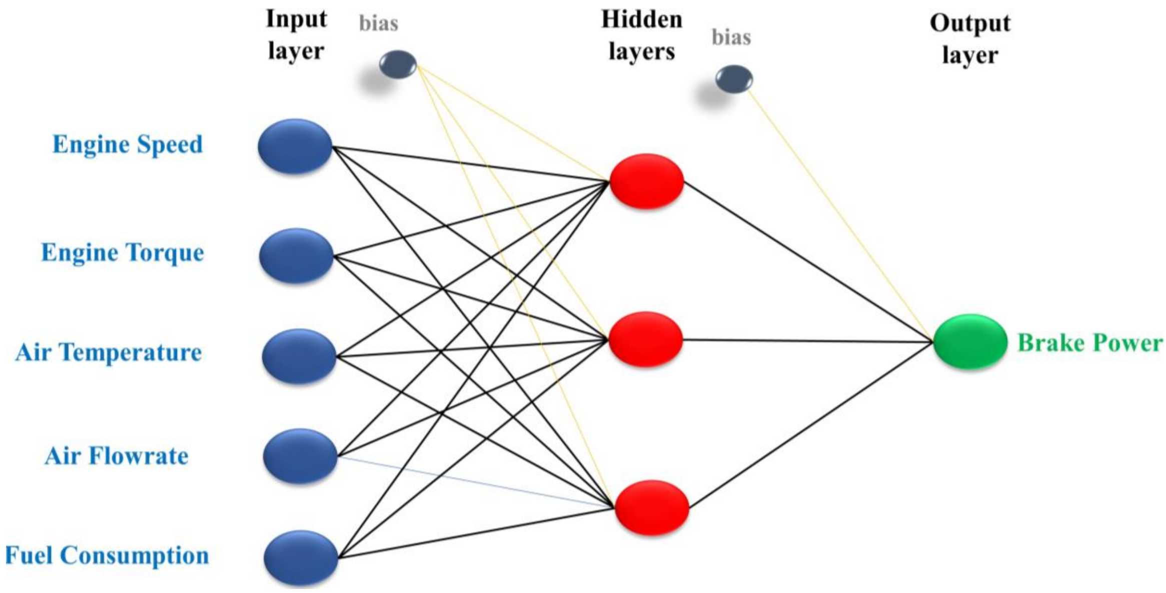

Fig. 3. ANN architecture

Table 3. Optimization of number of neurons in the hidden layer of ANN

Number of Neurons	BP		
	RMSE	R^2	MAE
3	0.0046	0.9952	1.51
4	0.0048	0.9949	1.57
5	0.0046	0.9952	1.52
6	0.0054	0.9935	1.79
7	0.0050	0.9945	1.67
8	0.0051	0.9941	1.72
9	0.0052	0.9940	1.75
10	0.0047	0.9951	1.55

2.3. SVM Model Setup

The support vector machine (SVM) model shown in Fig. 4 also depends on input and output parameters such as the ANN model [65]. Normalization is key for ensuring that every feature is equally considered in the model's decision-making process [66]. This is critical for increasing the model's sensitivity to data variations and increasing the speed of the training process. The SVM mode accepts the same input and output parameters (S, T, T_{air}, Q_{air}, and $\dot{m}$). The network's only output parameter is BP. To prepare for the modelling procedure, the dataset is subjected to normalization, and the input features are scaled to a homogeneous range, as shown in eq 9. We divided the dataset into three groups: training, validation, and testing. This section is critical for developing a strong model since it ensures comprehensive training and validation. It reduces overfitting and ensures that the model can predict any output from new data. Based on the training data, the SVM model uses a fit function to train the model by adjusting the internal settings of the SVR model. The model examines the data to understand the complicated relations between the input and output data, so this stage is considered the heart of the learning technique. In this model, the radial basis function (RBF) kernel is used to create the SVR model due to its effectiveness in handling nonlinear relationships, which contains hyperparameters such as regularization (C), epsilon (ε), and gamma (γ). These hyperparameters are commonly used in previous studies to adjust the complexity of the models according to the training data. The SVR model is customized with a set of selected data to obtain the highest performance for regression tasks. The kernel radial basis function is used because of its high performance in dealing with nonlinear data relationships. It allows the model to successfully translate input features into a higher-dimensional space for an accurate regression fit. The model is set on a regularization parameter, C = 100, to reduce training mistakes and produce an accurate fit to the training data [67], but at the cost of lower generalization. The epsilon parameter is ε = 0.01 to limit the tolerance range for acceptable predictions and guide the model through greater precision and sensitivity when fitting data [68]. The gamma parameter is set to ‘scale' to dynamically modify its value based on the variations in the data and changes in the curvature of the decision border and the model's ability to capture complicated patterns in the data. The selection of C and ε depends on the data characteristics and the desired precision of the model [69]. It is necessary that the model optimize between how precisely the model learns from training data and how successfully it applies that learning to new, previously unknown data. After training the model, the model is used to make predictions on a distinct set of data called the validation set. The flowchart for both models is shown in Fig. 5.

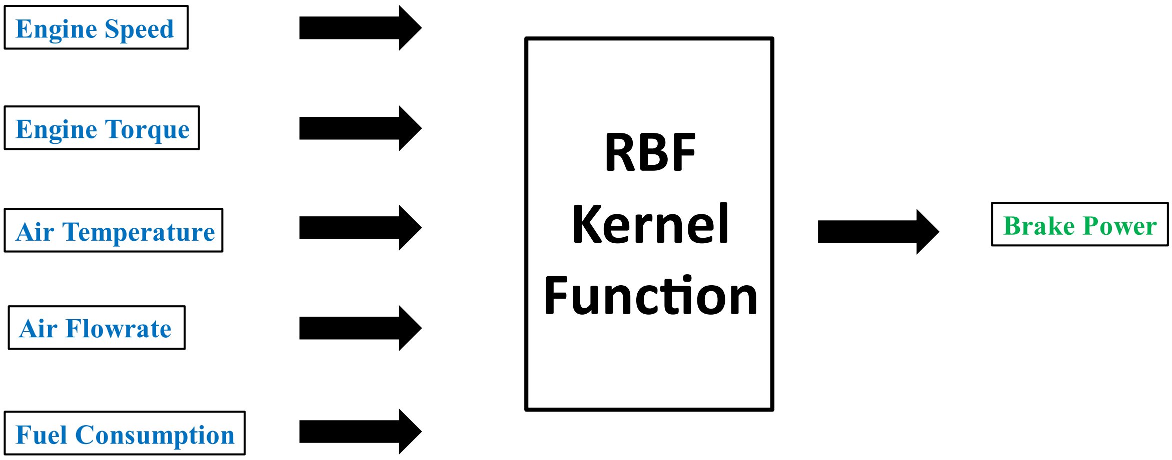

Fig. 4. SVM model architecture

3. Results and Discussion

The performances of the ANN and SVM models are compared in Fig. 5. The comparison between the two models is performed according to performance metrics such as the RMSE, EVS, and MAPE. Additionally, various comparative parameters are used, such as target vs. predicted values, regression plots, learning curves, histograms of prediction errors, residual plots, and cross-validation methods. Both models are known for their insightful information about how best to use them in the SI engine sector, especially in situations where accuracy and flexibility with large, complicated datasets are essential. The purpose of this comparison is to indicate the accuracy, complexity, and generalizability inherent in each strategy to help determine which model is best for a given task. According to the performance metrics, the complexity and practicality of the two models were as follows:

- Computational Requirements: In general, such ANN models require more computing due to their architecture and complicated training techniques that involve a large number of neurons and layers. SVM models, on the other hand, are computationally simpler when dealing with the linear kernel, making them more suitable for applications with high processing power.
- Training Time: The training process for the ANN model can take a long time, which requires the hardware support of a high caliber, particularly GPUs, and extended training time. SVM models are also, generally speaking, faster in terms of training, especially when using limited data or low-order kernels. This makes SVMs more suitable for use in situations where quick identification and an instant decision must be made.
- Ease of Use: ANN-based model deployment also proves to be a fairly challenging task, which demands the knowledge of neural architectures along with hyperparameters. On the other hand, SVM models are easier to apply due to their comparatively lower hyperparameters, especially for persons with less experience in ML.
- Real-World Applications: In practical conditions, this is dependent on the criteria of the specific problem and the particular strengths of either the ANN or the SVM. These depend on the size and complexity of the dataset as well as the relation patterns in the database, where in the cases of a large dataset that has complex and nonlinear relations, ANN models are preferred. Due to the ability of the networks to learn the many to many relations, several applications that require high accuracy have been proposed. On the other hand, SVM models are appealing because a simpler model that is easy to interpret and computationally less efficient is used. They are not prone to overfitting and are very easy to implement, especially when used in real-time applications and those that require fewer resources to operate.

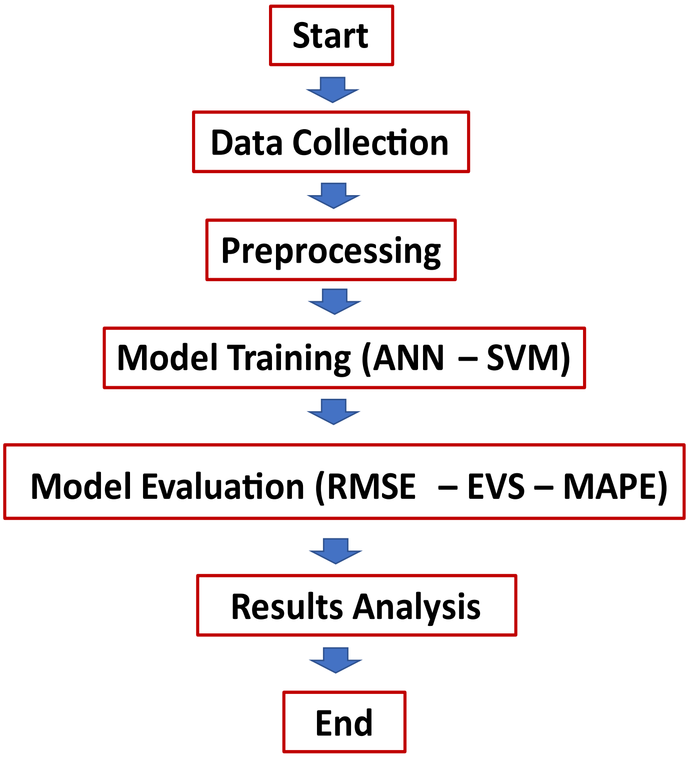

Fig. 5. Flowchart evaluation for ANN and SVM models

3.1. Performance Analysis and Evaluation

3.1.1. Performance Metrics for the SVM and ANN Models (RMSE, EVS, and MAPE)

According to the RMSE, EVS, and MAPE, the SVM and ANN are compared to evaluate the performance of each algorithm, as shown in Fig. 6. The SVM model has a lower RMSE (0.0044) than does the ANN model with a value of 0.0046, indicating that the SVM is accurate in predicting new data close to the target data and has a higher EVS (0.9953) compared to the ANN model with a value of 0.9952, indicating that the SVM indicates better model performance because it quantifies the percentage of the dependent variable's variation that can be predicted from the independent variables, suggesting that it is able to explain a greater percentage of the variance in the data, while the ANN is shown to have a lower value of MAPE (1.51%) compared to the SVM model with a value of 1.56%, suggesting that the SVM performs somewhat less well in terms of percentage error.

To support the performance claims, appropriate statistical tests of significance were conducted. As a consequence, the paired t test was used on the RMSE, EVS, and MAPE values achieved on cross-validated folds for both modes. The p values obtained from these tests are displayed in Table 4. For the RMSE, EVS, and MAPE, the p values are less than 0.05, which implies that the variations in the performances of the metrics between the SVM and ANN classifiers are not due to random sampling errors.

3.1.2. Target vs Predicted Values

The performances of the ANN and SVM models are compared in Fig. 7. The ANN model performs very well because its predicted values frequently match the target values, indicating a high degree of accuracy. This shows that the validation data fit the model quite well. It adjusts to changes in the target data accurately. Its capacity to represent unexpected rises and falls points to an accurate model with strong nonlinear variability handling capabilities. Across the whole dataset, the ANN predictions are in line with the goal values. This suggests that the model has probably successfully learned the underlying distribution of the data. Although it is not immediately apparent without testing on a different dataset, the ANN's close tracking of the target suggests that it may have learned specific characteristics and noise from the training data, which could be a risk for overfitting. The SVM model, on the other hand, shows more differences between the target and projected values and less accuracy. Predictions tend to be balanced, which may indicate more robust regularization or a more straightforward model. Since the SVM is less accurate at capturing extreme changes, it may be a sign of a simpler model or one that promotes simpler decision functions ahead of data variability fitting. The SVM predictions exhibit regular patterns of under- and overestimation at the peaks and troughs, respectively, in regard to the target values. Better generalization—which could lead to improved performance when faced with fresh, untested data-is suggested by the SVM's smoother prediction curve, albeit at the expense of short-term accuracy. The output of the ANN model has some fluctuations, as seen from sample points 2, 6, 7, and 8, where the predicted results are away from the target values, which may be a result of overtraining of the network and their abrupt changes. On the other hand, the SVM model, which is generally observed to yield smoothed response curves, is characterized by either underestimation or over-estimation in the region of the peak values, such as for samples 5, 6, and 7.

This indicates that even though the SVM model lacks flexibility enough to catch high variations, it is more generalized or more competent at providing stable predictions in other new untested data, proving more competency at the cost of high near-situation volatility [70].

Table 4. Comparison between SVM and ANN model in terms of statistical analysis of performance metrics

Measurement	Uncertainties
RMSE	0.02
EVS	0.03
MAPE	0.05

3.1.3. Regression Plots

We can evaluate the ANN and SVM models' performance by examining the degree to which the projected and actual outputs for the training, testing, and validation datasets agree in the regression plots shown in Fig. 8. The training, testing, and validation data predictions made by the ANN model are all very close to the regression line, showing a high degree of agreement between the expected and actual outputs. The data points and regression line almost exactly overlap, indicating a high prediction accuracy for the model. It appears that the ANN model generalizes well and does not overfit the training data based on the constant alignment observed across training, testing, and validation data. The ideal prediction line is shown in two models. The SVM model prediction values take the shape of a curve, which means that the data follow a nonlinear pattern in which a more complex function has been trained by the SVM model to match the data. The model works well on the data it was trained on, fitting almost perfectly. However, in regard to new data, it has not been seen before, as in testing; it does not do as well, especially for very high or very low values. This could mean that the model learned the training data too well and might not be good at predicting real-world situations. Moreover, it tends to become less accurate when the actual numbers start to increase.

3.1.4. Histograms of Prediction Errors

As shown in Fig. 9, the ANN model performed exceptionally well in terms of the prediction accuracy, exhibiting constant precision across the training, validation, and testing datasets. Its prediction errors were shown to be highly accurate, as they were tightly clustered around zero. The fact that all dataset error distribution patterns were comparable demonstrated the generalizability of the model. According to the reliability of the ANN model, the model was improved because it generated few outliers. The model is able to produce predictions that closely approximate the actual values, which is further supported by the prominent peaks in the density charts. The accuracy, generalizability, and resilience of the ANN model make it a good fit for practical applications. On the other hand, the SVM model yields greater variability in prediction errors with a larger spread, which suggests uneven predictive accuracy. A higher outlier frequency indicates that the model has a greater likelihood of larger errors. There were lower peaks in the density plots for the testing and validation data, which indicated some degree of confusion in the model's predictions. Due to these results, the SVM model performance and precision are still sufficient for some applications despite these constraints.

The training, validation, and testing data in each dataset are usually symmetrical around zero with slight fluctuations around it, as seen in the peaks of the curve, with at most a value of approximately 400 for testing. There are few training data points, approximately 350, and approximately 200 data points are used for the validation data, which indicates that the ANN model can be generalized. By comparing the distribution of the prediction errors, as illustrated by the histogram and the density plot, it is clearly observed that the SVM model has a wider variation in the errors. The number of training data examples is approximately 150, and the number of validation data examples is approximately 350, whereas the number of testing data examples is nearly 250. If the SVM is expanded across the entire dataset, it is clear that although the accuracy is high, the model is less consistent than the ANN model.

For both models, the prediction of stock prices is satisfactory; however, the ANN model has more tightly clustered prediction errors than the SVM, while the latter has recorded some accurate predictions but has greater variability in the errors observed [71].

3.1.5. Residual Plots

As shown in Fig. 10, in both graphs, the horizontal axis shows the values predicted by the models, while the vertical axis shows the residuals, which are the differences between the actual values and the predictions. The red dashed line at the zero mark on the vertical axis represents perfect predictions with no error. The ANN residuals indicated a dispersed range of values greater than that of the SVM model. This means that there is more uncertainty in the prediction values from the ANN model, but the residuals are mostly scattered near the zero line, with no clear pattern. There is one point that stands out far above the rest, indicating a significant prediction error for that particular data point, while the SVM's graph has its residuals more closely clustered around the zero line, suggesting smaller errors in prediction. However, there seems to be a slight trend where the residuals become more negative as the predicted values increase, indicating a consistent underestimation for larger values. For ANN model testing, smaller residuals approaching zero suggest good forecasting precision for values near 0.125 to 0.175. However, the residuals become more volatile, particularly at 0.225 to 0.250

and 0.275 to 0.300, which means that they have higher prediction errors than ideal random forests. The only noticeable discrepancy that was estimated to be beyond the forecasted value was 0.325, which can also indicate a miscalculation in terms of prediction. However, for the SVM model, the observed residuals are considerably close to zero for predicted values that are almost equal to zero (0.125 to 0.175), indicating good accuracy. The residuals for values of approximately 0.200 to 0.225 and 0.250 to 0.275 are, however, relatively accurate, with higher levels of variability. Residuals near 0.300 to 0. Moreover, 325 remained near zero, indicating good accuracy at the richness level of the predictions.

In summary, both the SVM and ANN models are good at performing all four functions of classification, but the residuals of the SVM are closer to zero most of the time and thus must be considered more consistent [72].

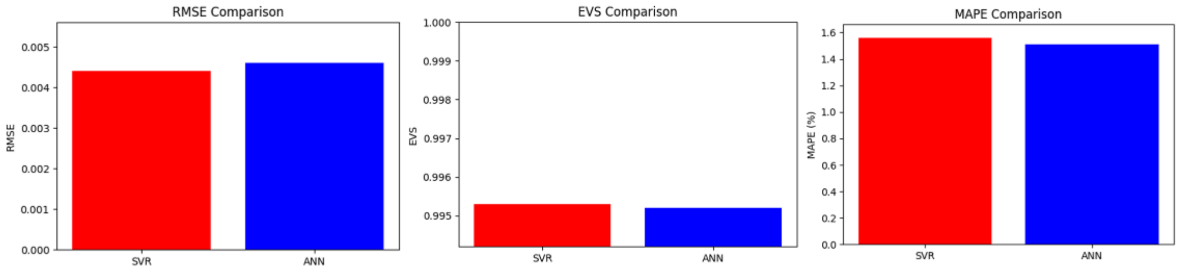

Fig. 6. Performance metrics comparison between SVM and ANN models, showcasing RMSE, EVS, and MAPE values

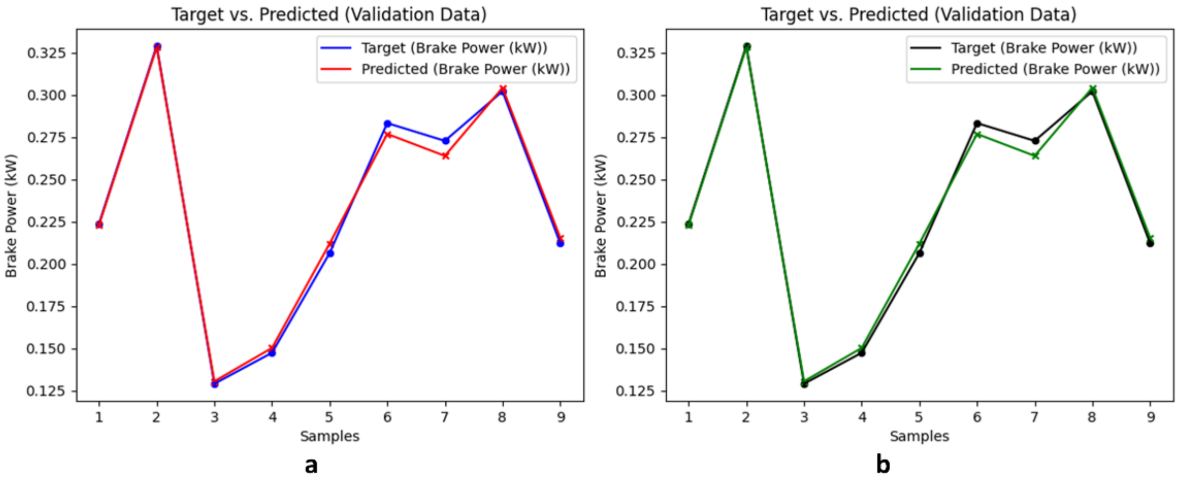

Fig. 7. Comparison of actual vs. predicted brake power (kW) using (a) ANN and (b) SVM models on validation data

3.2. Visualizing Data, Generalization, and Overfitting Protection Strategies

According to the ANN and SVM results, innovative visualizations were utilized to highlight the model's efficacy and the precise measures employed to reduce the risk of overfitting. By incorporating some approaches, it is necessary to conduct a transparent and rigorous analysis that reinforces the generalizability and dependability of the findings. The results of both algorithms are investigated using two methods: cross-validation and learning curve analysis.

3.2.1. Learning Curve

The learning curve in Fig. 11 of the ANN model begins with higher errors and gradually decreases as more data are added during the training phase. This behavior shows that feeding additional data makes the model work well in terms of the prediction of outputs, even though it may not perform well at first. After these high error values, the error decreases as the testing and validation error rates both significantly decrease to closely resemble the training error. ANNs have been shown to be able to reduce performance on unseen data, and their ability to efficiently adapt to new data has been demonstrated by their convergence at decreased error rates with increasing data volume, while the curve of the SVM model quickly decreases after a rapid decrease in the error rate is observed. This indicates that the model needs only a small quantity of data to comprehend the underlying data distribution. The SVM catches the important patterns early on, as demonstrated by the negligible change in error rates with more data. The model's capacity to be generalized from the beginning is demonstrated by the testing and validation error rates, which follow this trend of rapidly declining and then running parallel to the training error. The strong performance of the SVM is implied by the close association between these errors over a range of data sizes.

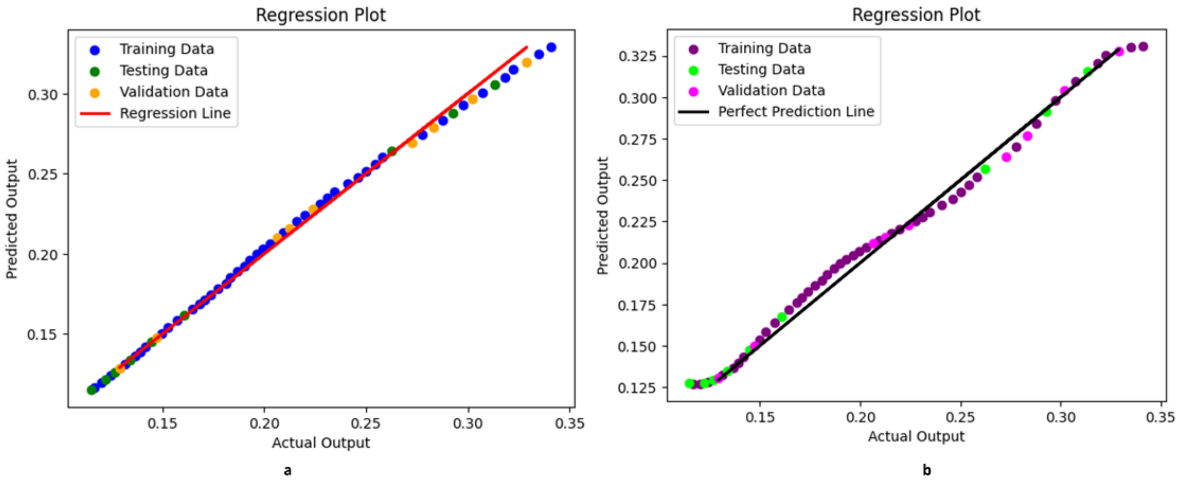

Fig. 8. Comparative regression analysis of (a) ANN and (b) SVM models, across training, testing, and validation datasets

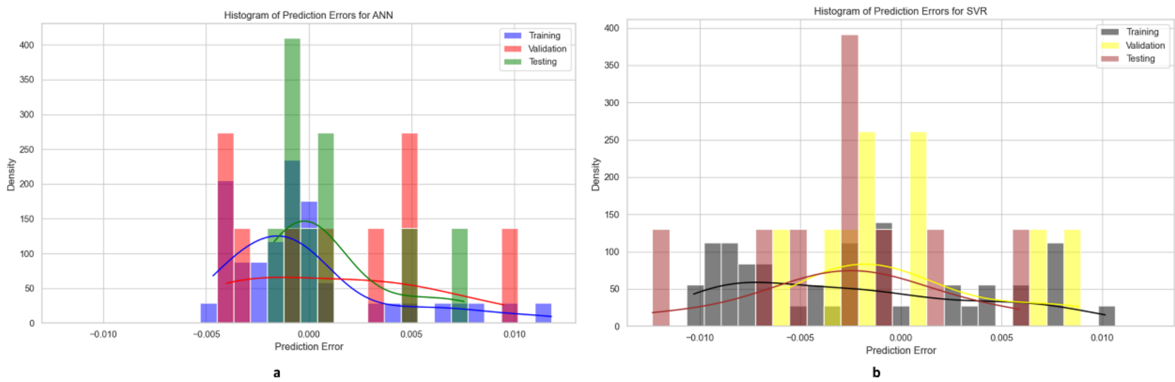

Fig. 9. Histograms of prediction errors for (a) ANN model and (b) SVM model

3.2.2. Cross-Validation

When applying the cross-validation technique to the two algorithms (ANN and SVM), as shown in Table 5, the R^2 values for the BP ANN range from approximately 0.9957 to 0.9985, with an average R^2 of 0.9966, while the R^2 values for the SVM show more variability, ranging from approximately 0.9846 to 0.9967, with an average R^2 of 0.9907. The high R^2 values obtained for the BP ANN and SVM across different folds of the dataset affirm their ability to generalize well to unseen data. The near R^2 values with high averages support the conclusion that BP does not overfit.

The results shown in Table 6 indicate that the ANN model achieved a mean RMSE of 0.0036 ± 0.0007, an EVS of 0.9970 ± 0.0008, and a MAPE of 1.23% ± 0.16%, whereas the SVM model achieved a mean RMSE of 0.0057 ± 0.0016, an EVS of 0.9949 ± 0.0033, and a MAPE of 2.54% ± 0.67%. Indeed, the Mohamed S. Hofny (Comparative Study of ANN and SVM Model Network Performance for Predicting Brake Power in SI Engines Using E15 Fuel) findings of the comparison

Vol. 4, No. 3, 2024, pp. 979-999

among the prediction models showed that the ANN model has lower RMSE and MAPE than the SVM model, as indicated in Fig. 12, which, as a result, shows higher accuracy and lower percentage error on the unseen data. It can be seen that the scores for EVS are again very similar to those based on the two models, although the ANN model has a slight advantage. It is now time to quantify the fact that the SVM model may be better at generalizing to other unseen instances through analysing other measures such as cross-validation techniques. The additional performance metrics and the statistical significance tests (paired t tests) also indicate the differences in the level of generalization of the models.

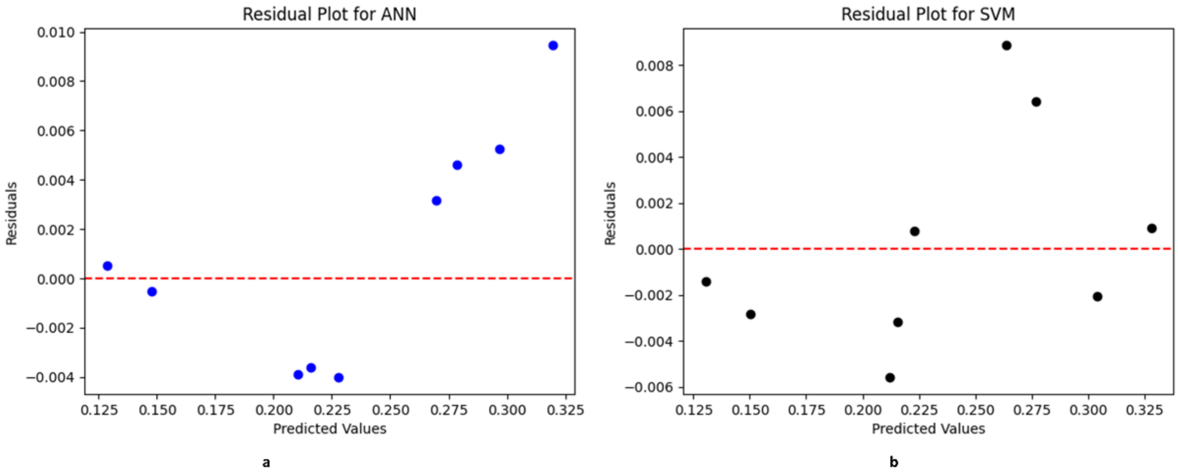

Fig. 10. Residual validation plots for both (a) ANN model (b) SVM model

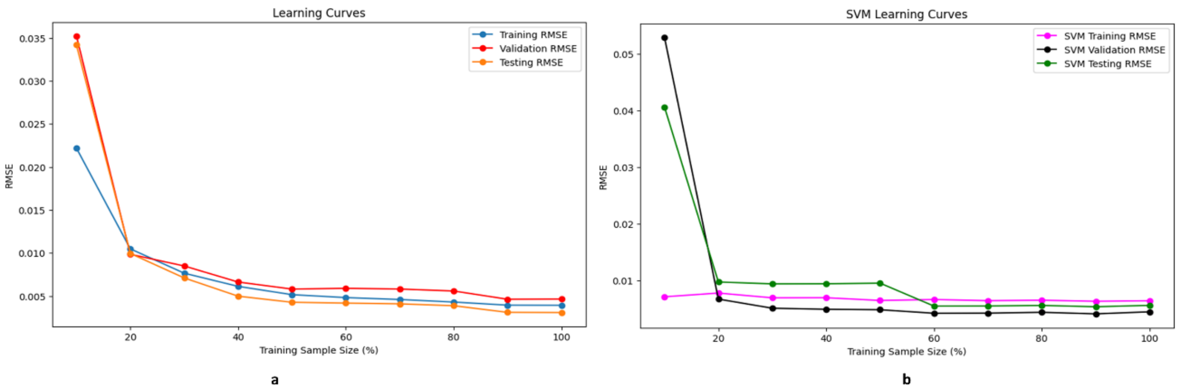

Fig. 11. Learning curves for both (a) ANN model and (b) SVM model

3.2.3. Robustness and Sensitivity Analysis

The actual sensitivity analysis was aimed at examining and comparing the stability and, at the same time, the sensitivity of the ANN and SVM models. Thus, this paper aims to quantify the behaviour of the models when the input parameters deviate from some predefined normal values, which is quite important for determining their stability in the case of their application to data containing significant noise. In the sensitivity analysis, the researcher entered random small noises into the input data and evaluated the performance measures (RMSE, EVS, MAPE) with the noisy data. To generate the perturbed dataset, a noise level of 5% of the standard deviation of the entire input data was added.

As seen from the findings in Table 7, both models slightly increase the RMSE and MAPE with distorted input data due to their responsiveness to input changes. However, the EVS values are not significantly affected and are constant, which ensures that the models still capture the variation in the data even with noise added. Thus, it can be concluded that both outcomes are rather stable to a certain extent, yet the ANN model is less sensitive to input variations than the SVM model. This robustness is important to guarantee that the models work under various scenarios and with noisy inputs, which increases the scope of their use in practical applications.

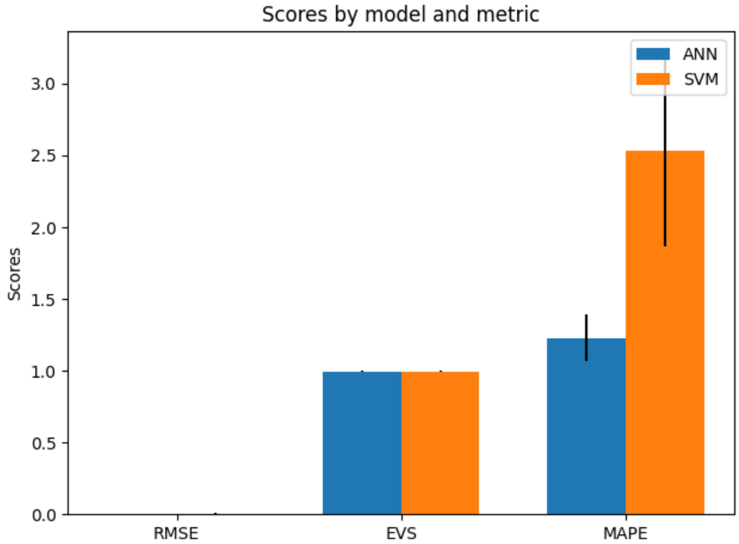

Fig. 12. Performance metrics of ANN and SVM algorithms

Table 5. K fold cross-validation R^2 scores for BP in ANN and SVM algorithms

ANN	SVM
[0.997792416162308, 0.9940754076638034, 0.9985444831809122, 0.9956917941883376, 0.9965629490071887, 0.996868747853836, 0.9961488436017671, 0.9966151990612169, 0.9961036682792997, 0.9971131870507753]	[0.9944836144167739, 0.9966706793643765, 0.9948164486902226, 0.9951218762877294, 0.9848213270539302, 0.984566612412749, 0.9912298668340159, 0.9868808614642891, 0.9852295773921526, 0.9935651976058324]
Average R^2: 0.9965516696049445	Average R^2: 0.9907386061522072

The performance differences between the ANN and SVM models are quite surprising and can directly impact real-life applications pertaining to the SI engine sector. The accuracy of the ANN model is higher; hence, it has smaller errors than the other models, making it more suitable for precise jobs such as adjusting the engine control unit and predictive maintenance, where accuracy in predicting the parameters of the engines is paramount for optimizing performance while inhibiting failure chances. On the other hand, the SVM model can be applied in situations with repetitiveness, real time, and high noise or variation in the data, for instance, onboard diagnostic and real-time monitoring techniques. However, for the same reasons, SVM has fewer equations, which translates to better interpretability, thus fitting the regulatory and safety analyses. Conclusively, limitations and assumptions that could affect the results presented in this comparative study of ANN and SVM models are as follows. The amount of data in the dataset, the quality of the data, the bias of the data, and the particular preprocessing steps influence the model and its overall performance. There is an importance of bringing tunable hyperparameters, and the specified configurations may not be the best universally. Moreover, the sensitivity analysis was conducted using simulated noise, the reality of which may not be quite similar to the one portrayed in the study. The SVM model has the drawback of assuming that accurate nonlinear separation of data may not always be achievable, whereas the ANN model is a disadvantage since the model is not easily understandable or explainable. When using the dataset to train the SVM model, it is assumed that the set can be effectively classified by a nonlinear function, an assumption that may sometimes be false, while the ANN model, as mentioned earlier, is not easily interpretable. Research could be conducted to determine how both ANNs and SVMs can be used together, as they both have advantages and disadvantages that may make them more helpful for use both in tandem. It may have more satisfactory results to apply it along with domain knowledge and a set of adaptive learning techniques to improve model interpretation and performance in practice in actual-time applications. The results could be further extended if the set of models were used for various types of engines and for various working conditions. According to these advances, the above investigations will surely advance SI engines and the application of machine learning techniques.

Table 6. Cross-validation performance metrics

Metric	ANN (Mean ± Std)	SVM (Mean ± Std)
RMSE	0.0036 ± 0.0007	0.0036 ± 0.0007
EVS	0.9970 ± 0.0008	0.9970 ± 0.0008
MAPE	1.23% ± 0.16%	2.54% ± 0.67%

Table 7. Sensitivity analysis performance metrics

Metric	ANN (Original)	ANN (Perturbed)	SVM (Original)	SVM (Perturbed)
RMSE	0.0036 ± 0.0007	0.0041 ± 0.0008	0.0036 ± 0.0007	0.0051 ± 0.0017
EVS	0.9970 ± 0.0008	0.9962 ± 0.0010	0.9970 ± 0.0008	0.9955 ± 0.0030
MAPE	1.23% ± 0.16%	1.51% ± 0.26%	2.54% ± 0.67%	2.20% ± 0.78%

4. Conclusion

According to the previous results, the ANN and SVM models are compared to indicate the performance of each model. A comparison between the two models is made through a comparison of the RMSE, EVS, and MAPE. Additionally, target vs. predicted values, regression plots, learning curves, histograms of prediction errors, and residual plots were generated. Cross-validation was used to determine whether there was overfitting, and robustness and sensitivity analyses were used to analyse the stability and sensitivity of the model to the input parameters. The support vector machine (SVM) model is indicated to have a lower RMSE (0.0044) and higher EVS (0.9953), while the ANN is shown to have a lower MAPE (1.51%). Additionally, the SVM model has less variance than the ANN model and thus has a lesser tendency to overfit and generalize better across datasets than does the ANN model. Robustness and sensitivity analysis also showed that both models are relatively insensitive to small changes in inputs, although the ANN model is somewhat more sensitive than the SVM model. Both models show some prediction errors, but there is no complex pattern in these errors that suggests a problem with the models' assumptions. The ANN seems to have a larger outlier, while the SVM shows a more consistent but possibly biased prediction. The residual and error analysis showed that large errors occasionally existed for the ANN model. The sample size is relatively small (60 samples) and can affect the accuracy, generalizability and robustness of each model, which can cause a certain extent of overlearning in the ANN model. Possible artefactual methods in the dataset could affect the generalizability of the results. Nevertheless, the comparative analysis of ANN and SVM models is useful for understanding the state of play in the SI engine sector. Therefore, it can be concluded that the use of the SVM model for SI engines is better because of its stable predictions, lower RMSE, higher EVS, and reliability. Future work will include enlarging the dataset, testing other methods and verifying the results in the real world. An ensemble of ANNs and SVMs can also be used to enhance the predictive accuracy of such models.

Author Contribution: All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] G. Joshi, A. Jain, S. R. Araveeti, S. Adhikari, H. Garg, and M. Bhandari, "FDA-Approved Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices: An Updated Landscape," *Electronics*, vol. 13, no. 3, p. 498, 2024, https://doi.org/10.3390/electronics13030498.

[2] T. K. F. Chiu, Z. Ahmad, M. Ismailov, and I. T. Sanusi, “What are artificial intelligence literacy and competency? A comprehensive framework to support them,” *Computers and Education Open*, vol. 6, p. 100171, 2024, https://doi.org/10.1016/j.caeo.2024.100171.

[3] Y. Shang, S. Zhou, D. Zhuang, J. Żywiołek, and H. Dincer, “The impact of artificial intelligence application on enterprise environmental performance: Evidence from microenterprises,” *Gondwana Research*, vol. 131, pp. 181-195, 2024, https://doi.org/10.1016/j.gr.2024.02.012.

[4] A. Dikshit *et al*., “Artificial Intelligence: A new era for spatial modelling and interpreting climate-induced hazard assessment,” *Geoscience Frontiers*, vol. 15, no. 4, p. 101815, 2024, https://doi.org/10.1016/j.gsf.2024.101815.

[5] M. Shoaei, Y. Noorollahi, A. Hajinezhad, and S. F. Moosavian, “A review of the applications of artificial intelligence in renewable energy systems: An approach-based study,” *Energy Conversion and Management*, vol. 306, p. 118207, 2024, https://doi.org/10.1016/j.enconman.2024.118207.

[6] B. Bala and S. Behal, “AI techniques for IoT-based DDoS attack detection: Taxonomies, comprehensive review and research challenges,” *Computer Science Review*, vol. 52, p. 100631, 2024, https://doi.org/10.1016/j.cosrev.2024.100631.

[7] G. Brydges, A. Uppal, and V. Gottumukkala, “Application of Machine Learning in Predicting Perioperative Outcomes in Patients with Cancer: A Narrative Review for Clinicians,” *Current Oncology*, vol. 31, no. 5, pp. 2727–2747, 2024, https://doi.org/10.3390/curroncol31050207.

[8] S. K. Khanal, A. Tarafdar, and S. You, “Artificial intelligence and machine learning for smart bioprocesses,” *Bioresource Technology*, vol. 375, p. 128826, 2023, https://doi.org/10.1016/j.biortech.2023.128826.

[9] S. A. Kalogirou, “Artificial intelligence for the modeling and control of combustion processes: a review,” *Progress in Energy and Combustion Science*, vol. 29, no. 6, pp. 515-566, 2003, https://doi.org/10.1016/S0360-1285(03)00058-3.

[10] L. Zhou, Y. Song, W. Ji, and H. Wei, “Machine learning for combustion,” *Energy and AI*, vol. 7, p. 100128, 2022, https://doi.org/10.1016/j.egyai.2021.100128.

[11] Ü. Ağbulut, A. E. Gürel, and Y. Biçen, “Prediction of daily global solar radiation using different machine learning algorithms: Evaluation and comparison,” *Renewable and Sustainable Energy Reviews*, vol. 135, p. 110114, 2021, https://doi.org/10.1016/j.rser.2020.110114.

[12] M. Bansal, A. Goyal, and A. Choudhary, “A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning,” *Decision Analytics Journal*, vol. 3, p. 100071, 2022, https://doi.org/10.1016/j.dajour.2022.100071.

[13] J. Fu *et al*., “Application of artificial neural network to forecast engine performance and emissions of a spark ignition engine,” *Applied Thermal Engineering*, vol. 201, p. 117749, 2022, https://doi.org/10.1016/j.applthermaleng.2021.117749.

[14] A. T. Hoang *et al*., “A review on application of artificial neural network (ANN) for performance and emission characteristics of diesel engine fueled with biodiesel-based fuels,” *Sustainable Energy Technologies and Assessments*, vol. 47, 2021, https://doi.org/10.1016/j.seta.2021.101416.

[15] C. Shi, C. Ji, H. Wang, S. Wang, J. Yang, and Y. Ge, “Comparative evaluation of intelligent regression algorithms for performance and emissions prediction of a hydrogen-enriched Wankel engine,” *Fuel*, vol. 290, p. 120005, 2021, https://doi.org/10.1016/j.fuel.2020.120005.

[16] M. H. Shahbaz and A. A. Amin, "Design of Active Fault Tolerant Control System for Air Fuel Ratio Control of Internal Combustion Engines Using Artificial Neural Networks," *IEEE Access*, vol. 9, pp. 46022-46032, 2021, https://doi.org/10.1109/ACCESS.2021.3068164.

[17] M. Munadi, M. Ariyanto, M. Muchammad, and J. D. Setiawan, “Optimal Engine Mapping Performances for Dual Spark-Plug Ignition Internal Combustion Engine Using Neural Network,” *Journal of Applied Engineering Science*, vol. 20, no. 1, pp. 195-205, 2022, https://doi.org/10.5937/jaes0-28542.

[18] G. J. Thompson, C. M. Atkinson, N. N. Clark, T. W. Long, and E. Hanzevack, “Technical Note: Neural network modelling of the emissions and performance of a heavy-duty diesel engine,” *Proceedings of the*

Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, vol. 214, no. 2, pp. 111-126, 2000, https://doi.org/10.1177/095440700021400201.

[19] L. Gosak, A. Svensek, M. Lorber, and G. Stiglic, "Artificial Intelligence Based Prediction of Diabetic Foot Risk in Patients with Diabetes: A Literature Review," *Applied Sciences*, vol. 13, no. 5, p. 2823, 2023, https://doi.org/10.3390/app13052823.

[20] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, H. Arshad, "State-of-the-art in artificial neural network applications: A survey," *Heliyon*, vol. 4, no. 11, p. E00938, 2018, https://doi.org/10.1016/j.heliyon.2018.e00938.

[21] A. Soltoggio, K. O. Stanley, and S. Risi, "Born to learn: The inspiration, progress, and future of evolved plastic artificial neural networks," *Neural Networks*, vol. 108, pp. 48-67, 2018, https://doi.org/10.1016/j.neunet.2018.07.013.

[22] A. B. Garg, P. Diwan, and M. Saxena, "Artificial Neural Networks based Methodologies for Optimization of Engine Operations," *International Journal of Scientific & Engineering Research*, vol. 3, no. 5, pp. 1-5, 2012, https://www.ijser.org/researchpaper/Artificial-Neural-Networks-based-Methodologies-for-Optimization-of-Engine-Operations.pdf.

[23] M. A. Hashim, M. H. Nasef, A. E. Kabeel, and N. M. Ghazaly, "Combustion fault detection technique of spark ignition engine based on wavelet packet transform and artificial neural network," *Alexandria Engineering Journal*, vol. 59, no. 5, pp. 3687-3697, 2020, https://doi.org/10.1016/j.aej.2020.06.023.

[24] S. Bhowmik, R. Panua, S. K. Ghosh, A. Paul, and D. Debroy, "Prediction of performance and exhaust emissions of diesel engine fuelled with adulterated diesel: An artificial neural network assisted fuzzy-based topology optimization," *Energy and Environment*, vol. 29, no. 8, pp. 1413-1437, 2018, https://doi.org/10.1177/0958305X18779576.

[25] H. Yaşar, G. Çağıl, O. Torkul, and M. Şişci, "Cylinder Pressure Prediction of An HCCI Engine Using Deep Learning," *Chinese Journal of Mechanical Engineering*, vol. 34, no. 1, p. 7, 2021, https://doi.org/10.1186/s10033-020-00525-4.

[26] N. Kuzhagaliyeva, A. Thabet, E. Singh, B. Ghanem, and S. M. Sarathy, "Using deep neural networks to diagnose engine pre-ignition," *Proceedings of the Combustion Institute*, vol. 38, no. 4, pp. 5915-5922, 2021, https://doi.org/10.1016/j.proci.2020.10.001.

[27] J. Xing, H. Wang, K. Luo, S. Wang, Y. Bai, and J. Fan, "Predictive single-step kinetic model of biomass devolatilization for CFD applications: A comparison study of empirical correlations (EC), artificial neural networks (ANN) and random forest (RF)," *Renewable Energy*, vol. 136, pp. 104-114, 2019, https://doi.org/10.1016/j.renene.2018.12.088.

[28] I. Omran *et al*., "Deep reinforcement learning implementation on IC engine idle speed control," *Ain Shams Engineering Journal*, vol. 15, no. 5, p. 102670, 2024, https://doi.org/10.1016/j.asej.2024.102670.

[29] K. Muralidharan and D. Vasudevan, "Applications of artificial neural networks in prediction of performance, emission and combustion characteristics of variable compression ratio engine fuelled with waste cooking oil biodiesel," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 37, pp. 915-928, 2015, https://doi.org/10.1007/s40430-014-0213-4.

[30] L. Shalahuddin, A. Suksmono, and Y. P. Sembiring, "Prediction of Internal Combustion Engine Performance Using Artificial Intelligence," *Majalah Ilmiah Pengkajian Industri*, vol. 14, no. 2, pp. 153-162, 2023, https://doi.org/10.29122/mipi.v14i2.4164.

[31] A. M. Deris, A. M. Zain, and R. Sallehuddin, "Overview of support vector machine in modeling machining performances," *Procedia Engineering*, vol. 24, pp. 308-312, 2011, https://doi.org/10.1016/j.proeng.2011.11.2647.

[32] R. Fabrice and N. Villa, "Support vector machine for functional data classification," *Neurocomputing*, vol. 69, no. 7-9, pp. 730-742, 2006, https://doi.org/10.1016/j.neucom.2005.12.010.

[33] M. H. Naveed *et al*., "Cellulosic biomass fermentation for biofuel production: Review of artificial intelligence approaches," *Renewable and Sustainable Energy Reviews*, vol. 189, p. 113906, 2024, https://doi.org/10.1016/j.rser.2023.113906.

[34] L. Deiss, A. J. Margenot, S. W. Culman, and M. S. Demyan, "Tuning support vector machines regression models improves prediction accuracy of soil properties in MIR spectroscopy," *Geoderma*, vol. 365, p. 114227, 2020, https://doi.org/10.1016/j.geoderma.2020.114227.

[35] B. Liu, J. Hu, F. Yan, R. F. Turkson, and F. Lin, "A novel optimal support vector machine ensemble model for NOX emissions prediction of a diesel engine," *Measurement: Journal of the International Measurement Confederation*, vol. 92, pp. 183-192, 2016, https://doi.org/10.1016/j.measurement.2016.06.015.

[36] J. I. Ghojel, "Review of the development and applications of the Wiebe function: A tribute to the contribution of Ivan Wiebe to engine research," *International Journal of Engine Research*, vol. 11, no. 4, pp. 297-312, 2010, https://doi.org/10.1243/14680874JER06510.

[37] H. Duan, Y. Huang, R. K. Mehra, P. Song, and F. Ma, "Study on influencing factors of prediction accuracy of support vector machine (SVM) model for NOx emission of a hydrogen enriched compressed natural gas engine," *Fuel*, vol. 234, pp. 954-964, 2018, https://doi.org/10.1016/j.fuel.2018.07.009.

[38] A. Norouzi, M. Aliramezani, and C. R. Koch, "A correlation-based model order reduction approach for a diesel engine NOx and brake mean effective pressure dynamic model using machine learning," *International Journal of Engine Research*, vol. 22, no. 8, pp. 2654-2672, 2021, https://doi.org/10.1177/1468087420936949.

[39] M. Aliramezani, C. R. Koch, and M. Shahbakhti, "Modeling, diagnostics, optimization, and control of internal combustion engines via modern machine learning techniques: A review and future directions," *Progress in Energy and Combustion Science*, vol. 88, p. 100967, 2022, https://doi.org/10.1016/j.pecs.2021.100967.

[40] R. Yang, T. Xie, and Z. Liu, "The Application of Machine Learning Methods to Predict the Power Output of Internal Combustion Engines," *Energies*, vol. 15, no. 9, p. 3242, 2022, https://doi.org/10.3390/en15093242.

[41] S. Zaichenko, S. Shevchuk, V. Opryshko, S. Pryadko and A. Halem, "Autonomous electric power source energy efficiency improvement by internal combustion engine gases distribution control," *2020 IEEE 7th International Conference on Energy Smart Systems (ESS)*, pp. 262-265, 2020, https://doi.org/10.1109/ESS50319.2020.9160085.

[42] N. M. Ghazaly, M. Abdel-Fattah, M. M. Makrahy, "Determination of Engine Misfire Location using Artificial Neural Networks," *International Journal of Vehicle Structures and Systems*, vol. 11, no. 4, 2019, https://doi.org/10.4273/ijvss.11.4.13.

[43] H. A. Hadi, A. Kassem, H. Amoud, S. Nadweh, N. M. Ghazaly, and N. Moubayed, "Using Active Filter Controlled by Imperialist Competitive Algorithm ICA for Harmonic Mitigation in Grid-Connected PV Systems," *International Journal of Robotics and Control Systems*, vol. 4, no. 2, pp. 581-605, 2024, https://doi.org/10.31763/ijrcs.v4i2.1365.

[44] A. N. Bhatt, N. Shrivastava, "Application of Artificial Neural Network for Internal Combustion Engines: A State of the Art Review," *Archives of Computational Methods in Engineering*, vol. 29, pp. 897-919, 2022, https://doi.org/10.1007/s11831-021-09596-5.

[45] J. Rezaei, M. Shahbakhti, B. Bahri, and A. A. Aziz, "Performance prediction of HCCI engines with oxygenated fuels using artificial neural networks," *Applied Energy*, vol. 138, pp. 460-473, 2015, https://doi.org/10.1016/j.apenergy.2014.10.088.

[46] S. Yilmaz and M. Z. Bilgin, "Modeling and simulation of injection control system on a four-stroke type diesel engine development platform using artificial neural networks," *Neural Computing and Applications*, vol. 22, pp. 1713-1725, 2013, https://doi.org/10.1007/s00521-012-1054-7.

[47] B. Ghobadian, H. Rahimi, A. M. Nikbakht, G. Najafi, and T. F. Yusaf, "Diesel engine performance and exhaust emission analysis using waste cooking biodiesel fuel with an artificial neural network," *Renewable Energy*, vol. 34, no. 4, pp. 976-982, 2009, https://doi.org/10.1016/j.renene.2008.08.008.

[48] G. Udaybhanu and V. Mahendra Reddy, "A hybrid GA-ANN and correlation approach to developing a laminar burning velocity prediction model for isooctane/blends-air mixtures," *Fuel*, vol. 360, p. 130594, 2024, https://doi.org/10.1016/j.fuel.2023.130594.

[49] M. S. Chowdhury, "Comparison of accuracy and reliability of random forest, support vector machine, artificial neural network and maximum likelihood method in land use/cover classification of urban setting," *Environmental Challenges*, vol. 14, p. 100800, 2024, https://doi.org/10.1016/j.envc.2023.100800.

[50] A. Kurani, P. Doshi, A. Vakharia, and M. Shah, "A Comprehensive Comparative Study of Artificial Neural Network (ANN) and Support Vector Machines (SVM) on Stock Forecasting," *Annals of Data Science*, vol. 10, pp. 183-208, 2023, https://doi.org/10.1007/s40745-021-00344-x.

[51] N. A. Almansour *et al*., "Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study," *Computers in Biology and Medicine*, vol. 109, pp. 101-111, 2019, https://doi.org/10.1016/j.compbiomed.2019.04.017.

[52] F. Li, W. Wang, J. Xu, J. Yi, and Q. Wang, "Comparative study on vulnerability assessment for urban buried gas pipeline network based on SVM and ANN methods," *Process Safety and Environmental Protection*, vol. 122, pp. 23-32, 2019, https://doi.org/10.1016/j.psep.2018.11.014.

[53] A. Sharma, R. Kumar, P. K. Varadwaj, A. Ahmad, and G. M. Ashraf, "A comparative study of support vector machine, artificial neural network and Bayesian classifier for mutagenicity prediction," *Interdisciplinary Sciences: Computational Life Sciences*, vol. 3, no. 3, pp. 232-239, 2011, https://doi.org/10.1007/s12539-011-0102-9.

[54] N. M. Ghazaly, A. O. Moaaz, M. M. Makrahy, M. A. Hashim, M. H. Nasef, "Prediction of misfire location for SI engine by unsupervised vibration algorithm," *Applied Acoustics*, vol. 192, p. 108726, https://doi.org/10.1016/j.apacoust.2022.108726.

[55] A. O. Hasan, H. Al-Rawashdeh, A. H. Al-Muhtaseb, A. Abu-jrai, R. Ahmad, and J. Zeaiter, "Impact of changing combustion chamber geometry on emissions, and combustion characteristics of a single cylinder SI (spark ignition) engine fueled with ethanol/gasoline blends," *Fuel*, vol. 231, pp. 197-203, 2018, https://doi.org/10.1016/j.fuel.2018.05.045.

[56] A. O. Hasan, H. Al-Rawashdeh, A. Abu-jrai, M. R. Gomaa, and F. Jamil, "Impact of variable compression ratios on engine performance and unregulated HC emitted from a research single cylinder engine fueled with commercial gasoline," *International Journal of Hydrogen Energy*, vol. 48, no. 68, pp. 26619-26628, 2023, https://doi.org/10.1016/j.ijhydene.2022.09.025.

[57] Y. Çay, A. Çiçek, F. Kara, and S. Sağiroğlu, "Prediction of engine performance for an alternative fuel using artificial neural network," *Applied Thermal Engineering*, vol. 37, pp. 217-225, 2012, https://doi.org/10.1016/j.applthermaleng.2011.11.019.

[58] I. Belakhdar, W. Kaaniche, R. Djmel and B. Ouni, "A comparison between ANN and SVM classifier for drowsiness detection based on single EEG channel," *2016 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pp. 443-446, 2016, https://doi.org/10.1109/ATSIP.2016.7523132.

[59] Ö. Çelik, "A Research on Machine Learning Methods and Its Applications," *Journal of Educational Technology and Online Learning*, vol. 1, no. 3, pp. 25-40, 2018, 2018, https://doi.org/10.31681/jetol.457046.

[60] M. J. Brown, L. A. Hutchinson, M. J. Rainbow, K. J. Deluzio, and A. R. De Asha, "A comparison of self-selected walking speeds and walking speed variability when data are collected during repeated discrete trials and during continuous walking," *Journal of Applied Biomechanics*, vol. 33, no. 5, pp. 384-387, 2017, https://doi.org/10.1123/jab.2016-0355.

[61] C. Sanjay and C. Jyothi, "A study of surface roughness in drilling using mathematical analysis and neural networks," *International Journal of Advanced Manufacturing Technology*, vol. 29, pp. 846-852, 2006, https://doi.org/10.1007/s00170-005-2538-8.

[62] W. K. Yap, T. Ho, and V. Karri, "Exhaust emissions control and engine parameters optimization using artificial neural network virtual sensors for a hydrogen-powered vehicle," *International Journal of Hydrogen Energy*, vol. 37, no. 10, pp. 8704-8715, 2012, https://doi.org/10.1016/j.ijhydene.2012.02.153.

[63] Xinghuo Yu, M. O. Efe and O. Kaynak, "A general backpropagation algorithm for feedforward neural networks learning," *IEEE Transactions on Neural Networks*, vol. 13, no. 1, pp. 251-254, 2002, https://doi.org/10.1109/72.977323.

[64] Z. Zhou, Q. Zhang, G. Lu, H. Wang, W. Zhang, and Y. Yu, “Adashift: Decorrelation and convergence of adaptive learning rate methods,” *International Conference on Learning Representations*, pp. 1-26, 2019, https://openreview.net/pdf/6e6f6adc0995fefe1616ec27b4c33639b4c7abdd.pdf.

[65] S. Gürgen, B. Ünver, and İ. Altın, “Prediction of cyclic variability in a diesel engine fueled with n-butanol and diesel fuel blends using artificial neural network,” *Renewable Energy*, vol. 117, pp. 538-544, 2018, https://doi.org/10.1016/j.renene.2017.10.101.

[66] S. Zhao, F. Blaabjerg and H. Wang, "An Overview of Artificial Intelligence Applications for Power Electronics," *IEEE Transactions on Power Electronics*, vol. 36, no. 4, pp. 4633-4658, 2021, https://doi.org/10.1109/TPEL.2020.3024914.

[67] H. Zhang, Z. Zhang, A. Odena, and H. Lee, “Consistency Regularization for Generative Adversarial Networks,” *Arxiv*, 2020, https://doi.org/10.48550/arXiv.1910.12027.

[68] C. Wang, P. T. Nguyen, Y. Wang, and T. Dang, “A Study of Performance Evaluation for Textile and Garment Enterprises,” *Processes*, vol. 10, no. 11, p. 2381, 2022, https://doi.org/10.3390/pr10112381.

[69] S. Bernard, L. Heutte, and S. Adam, “Influence of hyperparameters on random forest accuracy,” *Multiple Classifier Systems*, pp. 171-180, 2009, https://doi.org/10.1007/978-3-642-02326-2_18.

[70] T. Zhou, F. Wang, and Z. Yang, “Comparative analysis of ANN and SVM models combined with wavelet preprocess for groundwater depth prediction,” *Water*, vol. 9, no. 10, p. 781, 2017, https://doi.org/10.3390/w9100781.

[71] H. H. Shrimalve, S. A. Talekar, “Comparative Analysis of Stock Market Prediction System using SVM and ANN,” *International Journal of Computer Applications*, vol. 182, no. 1, pp. 59-64, 2018, https://doi.org/10.5120/ijca2018917444.

[72] M. Hasan, S. Ullah, M. J. Khan, and K. Khurshid, “Comparative analysis of SVM, ann and cnn for classifying vegetation species using hyperspectral thermal infrared data,” *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, pp. 1861-1868, 2019, https://doi.org/10.5194/isprs-archives-XLII-2-W13-1861-2019.