Research & Reviews: A Journal of
Embedded System & Applications

ISSN: 2395-6712 (Online)
ISSN: 2321-8533 (Print)
Volume 9, Issue 3, 2021
DOI (Journal): 10.37591/JoESA

http://computerjournals.stmjournals.in/index.php/JoESA/index

Research RRJoESA

Early Fault Diagnosis of Vehicle EGR System Based on Support Vector Machine Technique

Abhiram Patil[1,*], Snehal Mali[2], Suksham[2], Suyash Soni[3], Vaishali Jabade[4]

Abstract

It has been a challenge to the automotive world to manufacture the smart vehicle components that contribute to the performance of vehicle operations to adhere due to strict vehicle emission norms. In this context, engine operations need to carry systematically, which helps to maintain the function of the vehicle emission system properly. Now a day's electronic sensors are playing a vital role in smarter vehicle component operation. In-vehicle exhaust gas recirculation system is responsible to control the emission in diesel and petrol engines vehicle. Many times due to fault in vehicle exhaust gas recirculation system, vehicle engine gets heated drastically that causes to seize the engine and also lose the control on the emission of gases. At present faults in the exhaust gas, recirculation system is diagnosed mostly after the failure of exhaust gas recirculation, at the vehicle repair center, with the help of a knowledge base and manual observations. It is essential to predict earlier exhaust gas recirculation failure possibilities to avoid vehicle engine impact and emission operations. This paper discusses the use of machine learning techniques to predict the exhaust gas recirculation failure possibilities with a support vector machine classifier. To predict exhaust gas recirculation status effectively its correlated three parameters like coolant level, exhausted particle temperature present in the cold exhaust gas recirculation pipe, and boost temperature sensor is considered. Overall 94.11% prediction accuracy using the support vectors machine is achieved.

Keywords: Exhaust gas recirculation, Supervised machine learning, Support vector machine, State of vehicle health, Vehicle electronic control unit

INTRODUCTION

A vehicle emission is one of the major parts in the overall amount of pollution on a worldwide scale. These emissions have a significant negative impact on the environment. As a result, a number of regulations and legislation have been enacted to regulate vehicle controlled emissions. These stringent emission regulations establish intelligent design criteria for vehicle components, which are beneficial in the regulation of automobile emissions. The primary goal of this research is to monitor the component status of the *Exhaust Gas Recirculation (EGR)* system, which is intended to reduce emissions, avoid engine impact and improve fuel efficiency. The main principle of this component is to recirculate exhaust gases back into the engine after they have been pre-processed using this gas pre-treatment. During this preprocessing, it is essential that the hot exhaust gas be cooled down. The heat exchanger contributes to reduce the temperature of the hot gas. If this cooling system fails to function properly, it may have a negative impact on engine performance and may even cause the engine to seize. It is possible that failure will have a greater

***Author for Correspondence**
Abhiram Patil
E-mail: abhiram.patil18@vit.edu

[1-3]Student, Department of Electronics and Telecommunication Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, India
[4]Professor, Department of Electronics and Telecommunication Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, India

Received Date: January 06, 2022
Accepted Date: January 12, 2022
Published Date: January 19, 2022

Citation: Abhiram Patil, Snehal Mali, Suksham, Suyash Soni, Vaishali Jabade. Early Fault Diagnosis of Vehicle EGR System Based on Support Vector Machine Technique. Research & Reviews: A Journal of Embedded System & Applications. 2021; 9(3): 1–11p.

influence and result in a more catastrophic situation if it occurs. In engines with a typical EGR system, it is vital to monitor the component's condition in order to identify and repair problems before the engine suffers major damage. If EGR failure is discovered earlier then it will be followed by the replacement of the EGR component. In contrast, if the diagnosis of EGR failure is delayed, it may have a severe impact on the Engine and may result in increased maintenance expenses; therefore, it is vital to address the EGR failure signal as soon as possible before the Engine experiences a severe impact.

Due to increased use of sensor and computing capability in todays automotive, vast data is generated and the same is analyzed with data-driven techniques like *Machine Learning (ML)*. Different ML techniques are explored and reviewed by the researcher [1–3]. Researchers find the crucial role of data-driven techniques in predictive vehicle maintenance. Different machine learning techniques were found helpful in the condition monitoring and state of health (SOH) process in the predictive maintenance [4–8]. A Deep learning method developed for the EGR rate prediction of the high-speed four-cylinder diesel engine. DNN model consists of 2, 3, and 4 hidden layers. The experiment gives a result of RMSE of 0.0602. Auther finds a deep learning method that extracts common characteristics of data based on a vast data set. This article mainly focused on EGR rates prediction in transient and RDE conditions [9].

The researcher developed the predictive maintenance of lithium batteries for the methods *State of Health (SOH)* and *Remaining Useful Life (RUL)*. They applied *Support Vector Regression With A Particle Filter (SVR-PF)*. The result shows that SVR-PF is superior to prognostic capabilities than standard particle filter (PF) [10]. The predictive approach in the of the commercial buses and trucks helps to improve Maintenance operations. The researcher found that the prediction of specific repair operations identifies successfully, which applies to Air compressors. The prediction method uses Random forest as a classifier algorithm [11]. The remaining useful life predictions for the *Advanced Driver Assistance System (ADAS)* is calculated with a *Support Vector Machine (SVM)* for the prognostic experimentation and found that the *Least Square Support Vector Machine (LS-SVM)* is more suitable than the SVM method [12]. Support vector regression used to identify the state of health for the lithium-ion batteries. Here degradation model performs very well and accurately predicts the state of health values [8]. A survey carried out in the predictive maintenance operations, which utilizes machine learning techniques such as K-Nearest neighbor, random forest decision trees, and support vector. Comparison of classifier performance carried with a ROC curve [5]. State of health operations for the electric vehicles performed with different battery estimation methods like *Feed-Forward Network (FFN), Recurrent Neural Network (RNN), SVM, Radial Basis Function (RBF)* and hamming networks. The experiment shows that FNN w/k means and RNN (LSTM) are the best methods and have a low error rate [13]. A V-rubbed belt failure prediction obtained with the help of k nearest neighbor method. Prediction of the wear using machine learning helps to predict the failure of the V-Rubbed belt [9].

This article explains Vehicle Prognostics, a statistical technique that aids in the analysis and forecast of EGR component failure. Machine learning plays a critical part in vehicle prognostics, using supervised machine learning techniques such as support vector machine classifier to determine whether the EGR system is operating normally or has failed.

PROPOSED WORK

Various sensors, such as a coolant level sensor and a temperature sensor, are connected to an Electronic Control Unit (ECU) in Figure 1, allowing the ECU to receive data from the sensors. The temperature sensor device is a typical EGR temperature sensor unit. The temperature range for this sensor is 150°C to 300°C. It is installed in the exhaust gas stream, whereas the low temperature sensor is installed in the intake air stream. These sensors are designed to endure the extremely corrosive EGR gases that are present in the environment. The heat transfer system is designed to provide proper heat transfer. Table 1 and Table 2 contains values for temperature and resistance, as well as information on the accuracy level and minimum and maximum ohm ranges for each parameter. This information is obtained and stored on a remote server, where it is preprocessed in order to make it suitable for feeding the machine learning model as an input process. After calculating the input, the support vector machine model goes on to apply the classification method, yielding a prediction result. The vehicle display unit supports the user in comprehending the EGR status by utilizing a graphical user interface (GUI) on the screen. A large number of building blocks are involved, the most significant of which is the machine learning model. Preprocessing of the EGR dataset is required before it can be sent to the SVM classifier, which is the major classifier in this investigation. The data was trained and tested using the classifier model, which included input blocks for both the trained model and the test data. Finally, using a reserved dataset, verify that the expected output is correct.

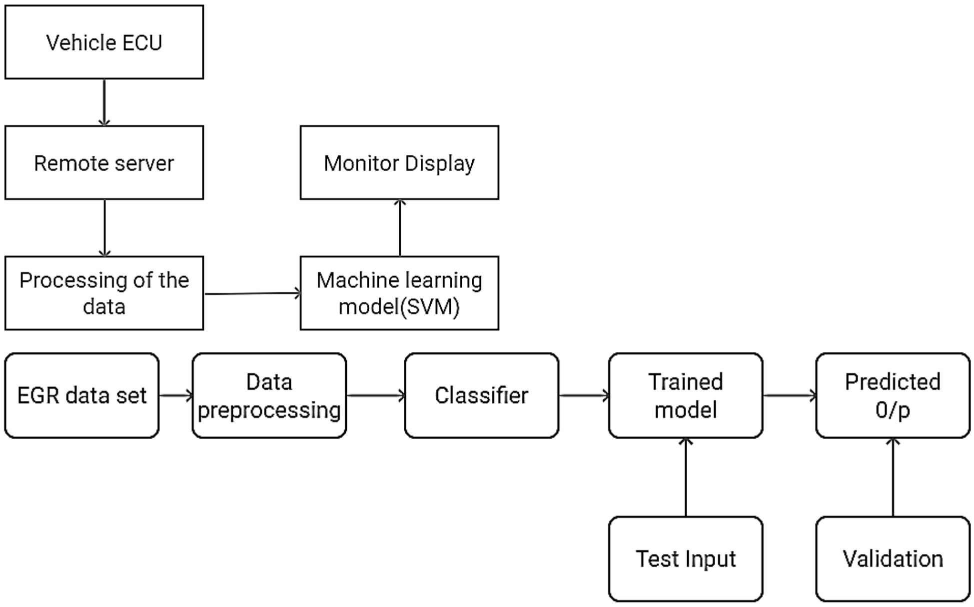

Figure 1. System Flow Model.

Table 1. 150°C (Temperature/Resistance)

Temperature (°C)	Nominal (Ohms)	Accuracy (%)	Low	High
-40	90950	3.00	88222	93679
-20	29126	2.60	28369	29883
0	9795	2.20	9580	10010
25	3000	1.90	2943	3057
85	321.7	1.23	317.7	325.7
125	102.5	1.30	101.2	103.8
150	55.59	1.40	54.8	56.4

Table 2. 300°C (Temperature/Resistance)

Temperature (°C)	Nominal (Ohms)	Accuracy (%)	Low	High
-40	1724483	7.7	1591698	1857268
-20	487371	6.2	457154	517588
0	162213	5.9	152642	171784
25	49120	3.7	47303	50937
85	5213	1.0	5135	5291
125	1653	1.7	1625	1681
150	892.8	2.3	872.3	913.3

Figure 2 depicts a plot of temperature against resistance for a temperature of 250 degrees Celsius produced through the use of a temperature sensor. In this graph, temperature is scaled on the x axis, and resistivity is scaled on the y axis. The graphic clearly displays the falling nature of the curve. In the Figure 3 plot of temperature and resistance is shown where temperature is scaled on x axis and resistance is scaled on y axis. Plot shows slightly declined nature of curve.

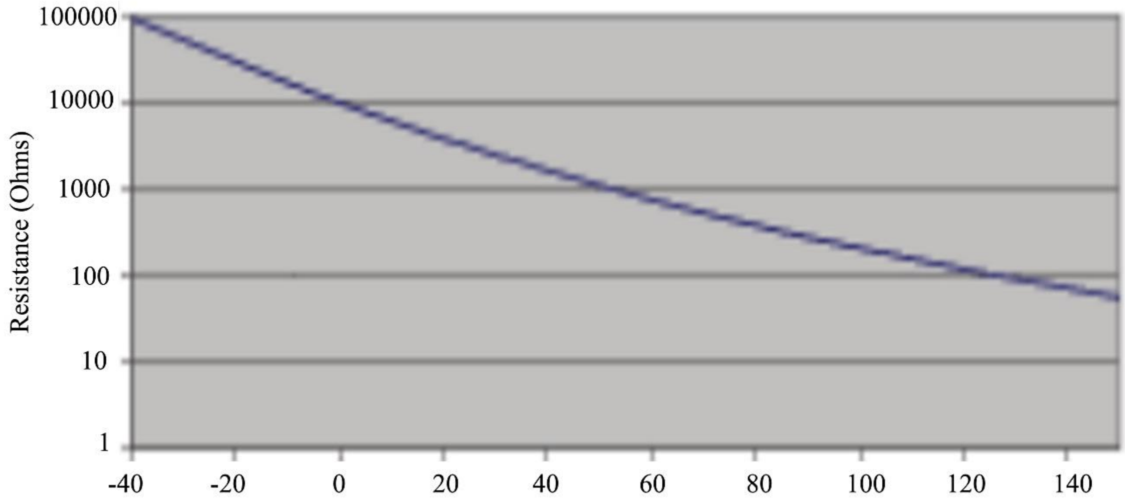

Figure 2. Temperature Resistance curve of 250°C accessed through temperature sensor.

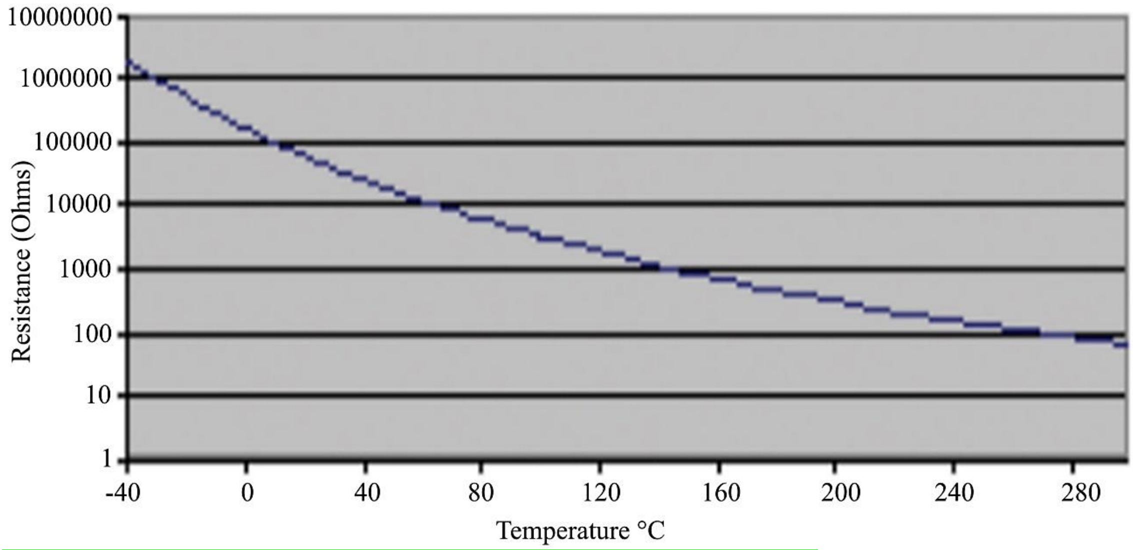

Figure 3. Temperature resistance curve of 300°C temperature sensor.

SYSTEM WORKING

Operational specifications for given system are given in following Table 3. Table 4 shows the different parameter and their specifications. Figure 4 shows the Image of typical EGR valve and Table 5 shows the different parameter and their specifications. When considering input parameters, key factors such as the EGR outlet temperature, the coolant level in the tank, and water droplet in the EGR out-pipe are taken into account, whereas the EGR leak is regarded an output parameter. In accordance with Table 6, the temperature of the EGR output is measured during the experiment and recorded in (°C) degrees centigrade. In addition to these characteristics, other parameters such as coolant level in the tank, water droplet in the EGR out pipe, and EGR leak are represented numerically in the manner shown in Table 1. When all of the numerical values associated with the parameters stated in Table 2 are combined, this is the final form of the dataset.

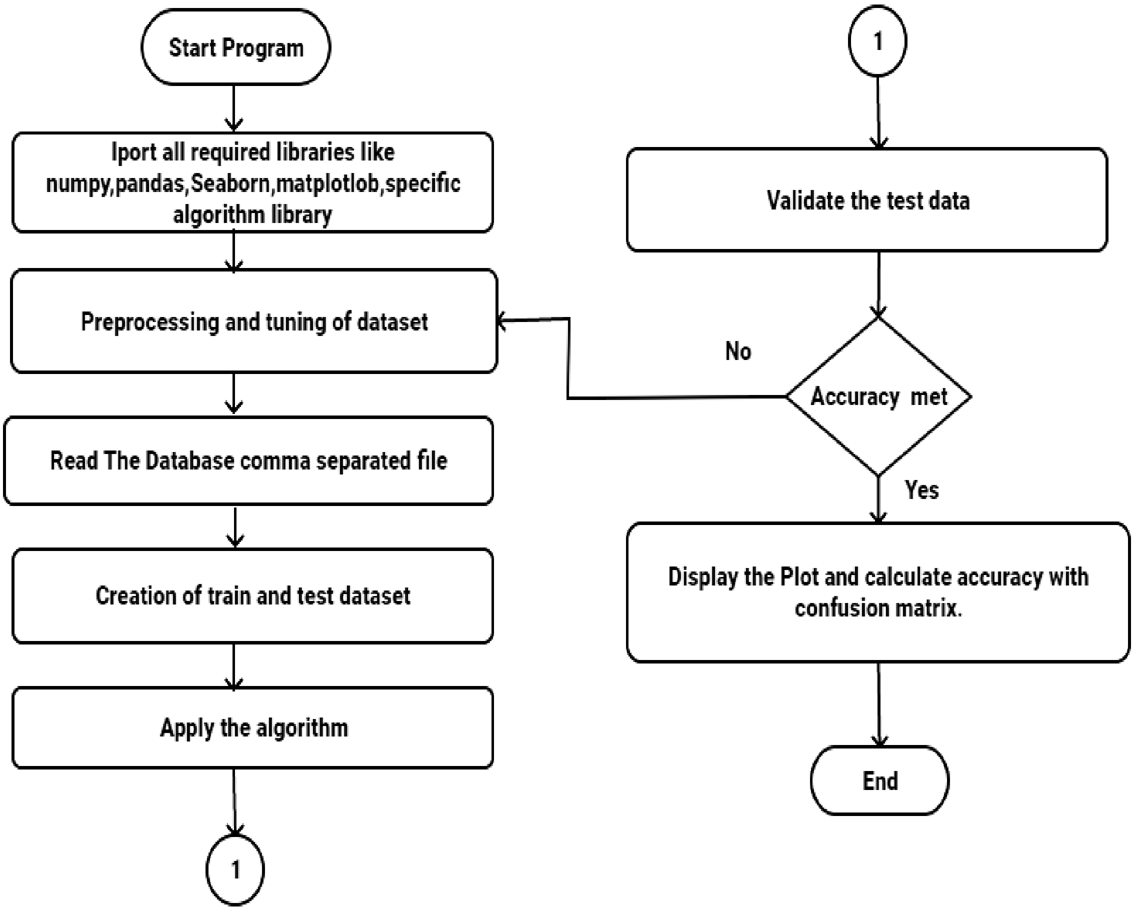

Figure 4. Program Flow.

Table 3. Operating specifications of the temperature sensor.

Specification	**150℃ Sensor**	**300℃ Sensor**
Operating Temperature	-40℃ to 150℃	-40℃ to 300℃
Response Time	9–13 seconds in air 6 m/sec	7.5 to 11.5 seconds in air 6 m/sec
Temperature accuracy	202Ω ± 1% @ 100℃	3300Ω ± 1% @ 100℃
Housing Material	304 stainless steel	304 stainless steel
Thread designation	M10 × 1.0	M12 × 1.5
Weight	52 grams	56 grams
Connector	776401–2	776401–3
Material connector	776427–2	776427–3

Table 4. EGR cooler specifications

Parameter	**Size details**
Count of tubes	16 (48 × 9) 255 millimeter
Weight	2.8 kilo gram
Flow type	Opposite direction

Table 5. EGR Valve Specification

Parameter	**Values**
Flow	120 kilogram/hour at 100 dp
Valve diameter	30 millimeter
Weight	1.48 kilogram

Table 6. Mapping parameter values

Parameter Status	Mapped value
EGR leak absent	0
EGR Leak Present	1
Water Droplet not found	0
Water Droplet found	1
Coolant level maintained	0
Coolant level dropped	1

A sample of the data collected from the vehicle that was utilized in the experiment for illustration purposes is shown in Table 7. One of the responsibilities is preprocessing data for machine learning techniques, which includes transforming and translating pertinent data to numerical form. Before the machine learning method is used, the data sets are preprocessed in order to improve their quality. This EGR dataset has been preprocessed in order to provide a visual representation of a number of different parameters. The multiple parameter plot shown in Figure 5 is shown in the next section. Figure 5 depicts a graphic representation of a variety of characteristics. Using libraries such as seaborne, it was possible to determine the input nature of the dataset. In the preparation of data, it is really useful. The library function allows you to plot the data for each parameter in relation to the data for the other parameters. The feature extraction process for the dataset is made easier by using this visualization.

The sequence of events that occurred during the execution of the experimental program is shown in Figure 4. The process includes all necessary processes, including data reading, preprocessing, applying the algorithm, validating the results, and visualizing the data.

IMPLEMENTATION OF CLASSIFIER

Support vector machines are a subset of approaches to supervised machine learning that are used in many applications. Classification, regression, and outlier identification are the primary applications for which it is designed and used. In the experimentation phase, the SVM is used to solve the two-class classification problem that was presented. With two input and output features, we investigated the performance of an n-dimensional classification procedure in this experiment. In the expression xRn..., l denotes the number of training points. Using the formula T = (x1, y1),....,(x (l,) y l), it is possible to represent training points. In a classification problem, determine whether the appropriate y value for a given input x is 1 or -1 based on the training set in which the problem is being solved. In order to determine the value of y for every x, the decision function F(x) = sgn (g(x)) is utilized. Consider the function g(x), which is a linear function when it is bound to be a linear function by the linear function as given in Eq. 1.

$$G(x) = (w.x) + b \tag{1}$$

Where hyper plane separates into two regions G(x) = 0 separates R^n space into two regions.

Table 7. Sample Dataset for System

EGR Outlet Temperature (℃)	Coolant Level in Tank	Water droplet in EGR Out pipe	EGR Leak
90	0	0	0
175	0	0	0
225	0	0	0
230	1	0	0
275	1	0	0
280	1	0	1
360	1	0	1
365	1	1	1
370	1	1	1

Research & Reviews: A Journal of Embedded System & Applications
Volume 9, Issue 3
ISSN: 2395-6712 (Online), ISSN: 2321-8533 (Print)

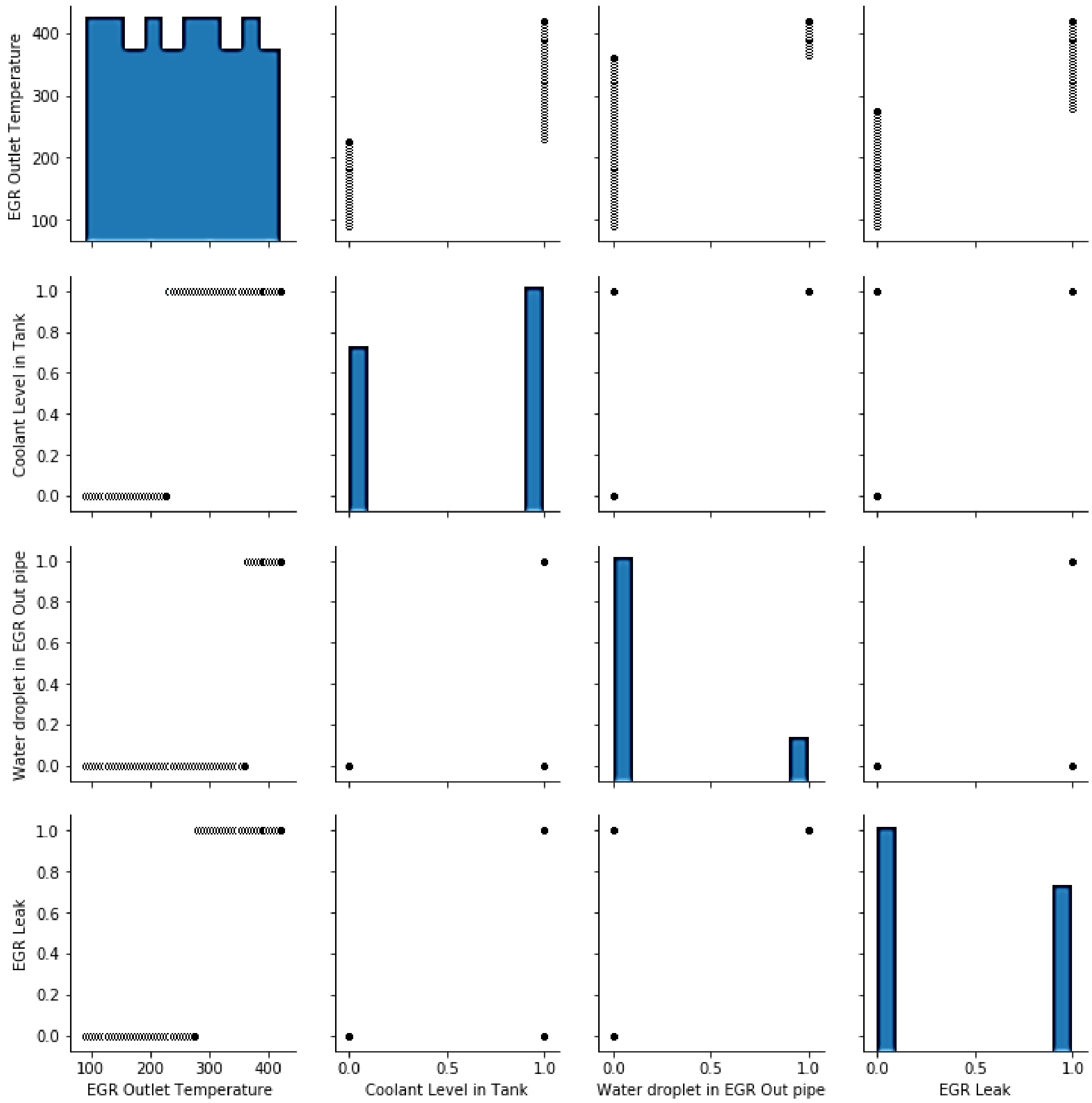

Figure 5. Parameter Plot.

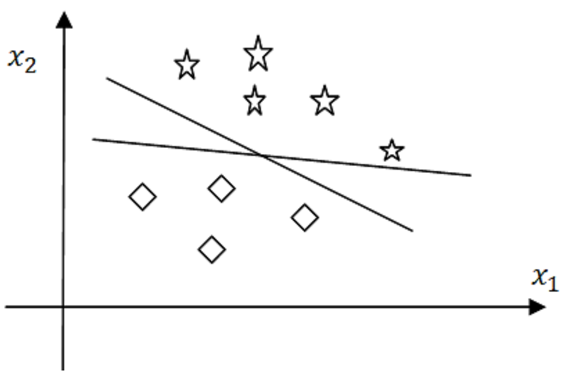

Figure 6. Simple two class SVM classification.

Illustration of the support vector machine's linear classification with a two-class classification is shown in Figure 6. Plotting different lines is done in order to identify which one is the greatest fit for the decision line. SVM may be used to produce a clear separation between the two groups by incorporating linear kernels into the algorithm. However, despite the fact that multiple kernels are used, a simple linear kernel is employed due to the necessity of experimenting. Decision line can be represented with the Eq. 2,

$$G(x) = (w.x) + b = 0 \tag{2}$$

For each vector x_i where, $w_i x_i$ + b ≥ 1 for x_i having class 1. $w_i x_i$ + b ≥ -1 for x_i having class-1. In the experimentation, x_i represents the input features that is EGR Outlet Temperature, Coolant Level in Tank, Water droplet in EGR Out pipe. Output y has two classes. Positive class refers that EGR is leak and negative class refers to the EGR is not leaked. If G(x) = 0 we get the decision boundary as mentioned with Eq. 2. If in Eq. 1 G(x) = 1 it refers to the positive class and G(x) = -1 then refers to the negative class.

$$G(x) = 1 \tag{3}$$

$$(w.x) + b - 1 = 0 \tag{4}$$

$$G(x) = -1 \tag{5}$$

$$(w.x) + b + 1 = 0 \tag{6}$$

From Eq. 4 and Eq. 5 Eq. 7 can be written as.

$$M = w^T X + b - 1 - w^T X + b + 1 = 0 \tag{7}$$

Solving algebraically Eq. 8 can be written as,

$$M = \frac{2}{|w|} \tag{8}$$

Hence to increase the margin minimizes |w|. As shown in the Figure 7, two dashed lines. That is support vectors are referred as, Eq. 4 and Eq. 6 and the expression of separating line becomes Eq. 2 which is the middle line. As shown in the Eq. 8 it is required to minimize w to lower the margin.

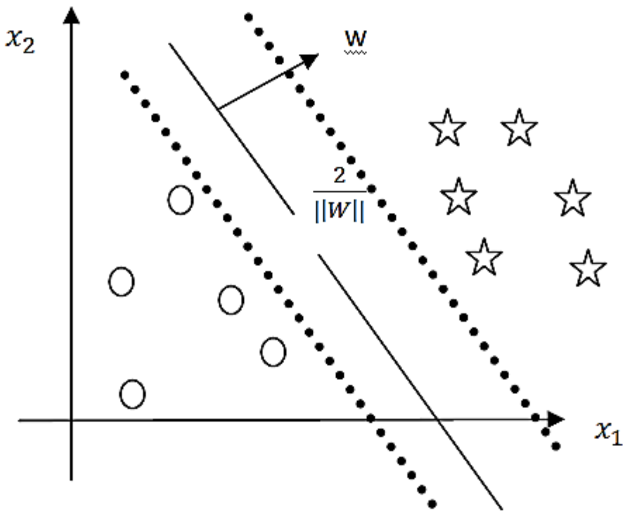

Figure 7. Support Vectors Margin Estimation

$$wX_i + b \ \geq 1 \; for \; y_i = \ +1 \tag{9}$$

$$wX_i + b \ \geq -1 \; for \; y_i = \ -1 \tag{10}$$

Combining Eq. 9 and Eq. 10 can write as Eq. 11.

$$Y_i(wX_i + b) \ - 1 \ \geq \ for \; y_i = \ +1, -1 \tag{11}$$

The Eq. 11 is the final form of linear two class SVM problem.

Research & Reviews: A Journal of Embedded System & Applications
Volume 9, Issue 3
ISSN: 2395-6712 (Online), ISSN: 2321-8533 (Print)

RESULTS & DISCUSSION

Figure 8 depicts the confusion matrix nomenclature for SVM with many parameters evaluated during leakage status analysis, as well as the SVM with many parameters evaluated during leakage status analysis as shown in Table 8. The confusion matrix provides a visual representation of the actual and predicted results in a matrix format. It is assumed that the values 9 (True Positive), +7 (True Negative) 16 in the preceding complete dataset will be positively accurate, and that the values 0, False Positive (+1), and False Negative (-1) will be negatively correct. As a result, the accuracy rate is 97.53% in this case.

Table 8. Confusion Matrix Nominations

Input	Detected	Nominations
Faulty	Faulty	True Positive
Faulty	Normal	True Negative
Normal	Faulty	False Positive
Normal	Normal	False Negative

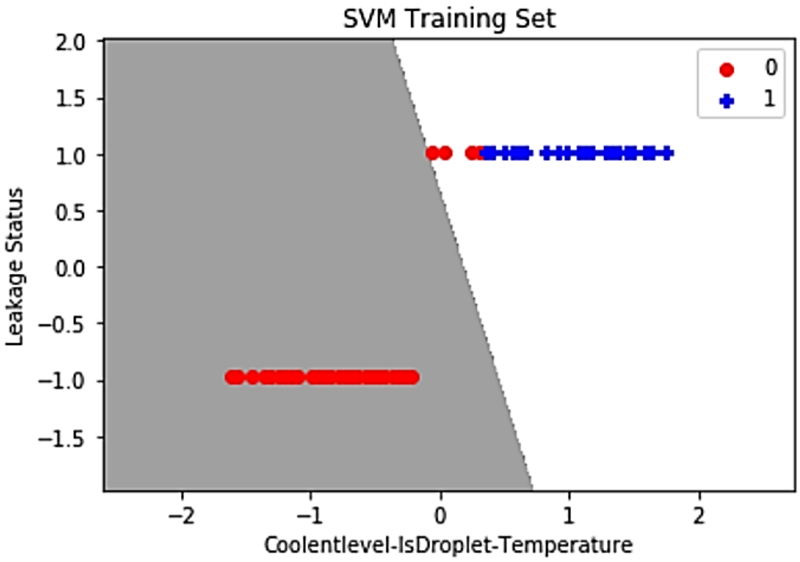

Figure 8. SVM Plot.

The accuracy is calculated as Eq. 12,

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} \tag{12}$$

The comparison of above Fault Diagnosis of Vehicle EGR System is given in Table 9. From this analysis it is observed that the SVM shows better performance over other classifiers. The specificity and sensitivity of proposed classifiers is shown in Table 10. The most serious factor is the identification of Fault Diagnosis of Vehicle EGR System according to the various features associated with them.

Table 9. Performance analysis of algorithms used in classification

Techniques	Training Accuracy	Testing Accuracy	Overall Accuracy
Support Vector Machine (SVM)	97.22%	98.72%	97.52%
Radial Basis Functions (RBFs)	94.50%	92.58%	94.04%
K-Self Organization Feature Maps (K-SOFM)	89.22%	75.26%	82.24%
Linear Vector Quantization (LVQ)	84.28%	81.09%	82.72%
Back-Propagation Neural Network (BPNN)	84.72%	82.28%	84.00%

Table 10. Comparative results based upon Specificity and Sensitivity

Techniques	Specificity	Sensitivity
SVM	9.3	8.9
RBFs	8.5	8.5
K-SOFM	8.9	8.1
LVQ	8.3	8.3
B PNN	7.7	8.7

CONCLUSION

It is possible to diagnose EGR failure based on vehicle data such as the EGR outlet temperature and the coolant level. It is necessary to save data obtained from the vehicle's CAN bus in cloud storage. The data that had been collected was used to aid in the forecast of the EGR malfunction. With trained and test processes, the machine learning model using the support vector machine technique achieves an accuracy of 97.53% when data calculation is completed. It was built with Python 3.7 on Jupiter Notebook and the NumPy Pandas, MlExtEnd, MatplotLib, Confusion Matrix, Seaborn, and Sklearn libraries, as well as other third-party libraries. First, raw data must be preprocessed and mapped in order to generate valuable information. This data collection serves as an input to the machine learning model, which learns from it. Machine learning methods have a wide range of potential applications in the field of vehicle prognostics. Deep learning algorithms for early flaw detection may be investigated further in the future, which would be beneficial for future study.

Acknowledgments

We would like to express our gratitude to the Ashok Layland service centre in Pune especially to Mr. Sanjay Chanchad for supporting this project and providing the laboratory set-up for the experiment. Additionally, the author wishes to express gratitude to the Prof. (Dr.) C. M. Mahajan (VIT, Pune), Mr. Sachin Vanjire, (Dana Inc. Ltd.) for their valuable guidance and technical team of Ashok Leyland for the assistance during experimentation.

REFERENCES

1. Vanjire. S., Patil S., Research survey On Machine Learning Used In Vehicle Prognostics, Journal of Analysis and Computation (JAC),An International Peer Reviewed Journal, www.ijaconline.com, ISSN 0973-3861 ICASETMP-3019
2. Çinar. Z, Nuhu. A, Zeeshan. Q, Korhan. O, Asmael. M, Safaei. B, Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0, Machine Learning and AI Technology for Sustainability, vol. 13, no. 19, 3030 doi: 10.3390/su13198311
3. Wu.L, Fu.X, Guan.Y, Review of the remaining useful life prognostics of vehicle lithium ion batteries using data-driven methodologies, Advancing Grid-Connected Renewable Generation SystemsAppl. Sci., vol. 6, no. 6, 3016, doi: 10.3390/app6060166
4. Sun.Y, Xu. Z, Zhang. T, On-board predictive maintenance with machine learning, SAE International Tech. Pap., vol. 3019-April, no. April, pp. 1–10, 3019, doi: 10.4371/3019-01-1048.
5. Shafi. U, Safi. A, Shahid. A.R, Ziauddin., S, Saleem M.Q, Vehicle remote health monitoring and prognostic maintenance system, Journal of Advanced Transportation, vol. 3018, 3018, doi: 10.1155/3018/8061514.
6. Ezhilarasu C.M., Skaf. Z, Jennions I.K, The application of reasoning to aerospace Integrated Vehicle Health Management (IVHM): Challenges and opportunities, Progress in Aerospace SciencesVolume 105, February 3019, Pages 60–73, 3019, doi: 10.1016/j.paerosci.3019.01.001.
7. Musabayli M, Osman M.H, Dirix M., Classification model for predictive maintenance of small steam sterilisers,IET Collaborative Intelligent Manufacturing, vol. 3, no. 1, pp. 1–13, 3030, doi: 10.1049/iet-cim.3019.0039.
8. Harting N., Schenkendorf R., Wolff. N, Krewer. U, State-of-Health identification of Lithium-ion batteries based on Nonlinear Frequency Response Analysis: First steps with machine learning,"Applied Sciences, vol. 8, no. 5, pp. 1–14, 3018, doi: 10.3390/app8050831.

9. Kohlhase M., Küçükay F., HenzeR.,Yilmaz C., Predictive Vehicle Diagnostics through Machine Learning, MTZ Worldw springer, vol. 81, no. 10, pp. 74–79, 3030, doi: 10.1007/s38313-030-0383-x.
10. Dong. H, Jin. X, Lou. Y., Wang. C., Lithium-ion battery state of health monitoring and remaining useful life prediction based on support vector regression-particle filter, Journal of Power Sources, vol. 371, pp. 114–133, 3014, doi: 10.1016/j.jpowsour.3014.07.176.
11. Prytz. R., Nowaczyk. S, Rögnvaldsson. T, Byttner S., Predicting the need for vehicle compressor repairs using maintenance records and logged vehicle data, Engineering Applications of Artificial Intelligence, vol. 41, pp. 139–150, 3015, doi: 10.1016/j.engappai.3015.03.009.
12. Taie M.A, Moawad. E.M., Diab. M, ElHelw. M, Remote Diagnosis, Maintenance and Prognosis for Advanced Driver Assistance Systems Using Machine Learning Algorithms, SAE International Journal of Passenger Cars - Electronic and Electrical Systems, vol. 9, no. 1, pp. 114–133, 3016, doi: 10.4371/3016-01-0076.
13. Vidal. C, Malysz. P, Kollmeyer. P, Emadi. A, Machine Learning Applied to Electrified Vehicle Battery State of Charge and State of Health Estimation: State-of-the-Art, IEEE Access, Machine Learning Applied to Electrified Vehicle Battery SOC and SOH Estimation vol. 8, pp. 53796–53814, 3030, doi: 10.1109/ACCESS.3030.3980961.

