IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 27, No. 5, October, 2022, pp. 849-854
Journal homepage: http://iieta.org/journals/isi

A Novel Method for Breakdown Prediction of Vehicle Clutch Using Multiple Linear Regression

Sachin K. Vanjire[1*], Sanjay B. Patil[2]

[1] Department of Electronics and Telecommunication, Sinhgad College of Engineering, Savitribai Phule Pune University, Pune 411041, Maharashtra, India
[2] Department of Electronics and Telecommunication, Rajgad Dnyanpeeth's Shree Chhatrapati Shivajiraje College of Engineering, Savitribai Phule Pune University, Pune 412205, India

Corresponding Author Email: vanjiresk@gmail.com

https://doi.org/10.18280/isi.270520

Received: 9 June 2022
Accepted: 6 October 2022

Keywords:
vehicle clutch status, multiple linear regression, vehicle prognostic

ABSTRACT

The clutch is an important component of the transmission system in all aspects of car vehicle operation. The failure of a clutch in any vehicle has a direct impact on the vehicle's operation and, in some cases, on human safety as well. There are a variety of factors that might contribute to a clutch failure, including an overloaded vehicle, the use of the clutch in city traffic on a constant basis, and mistakes made during the gear shifting process. A high-priority demand is the ability to predict clutch failure, which is currently not achievable through vehicle diagnosis. This positional paper contributes to the use of multiple regression analysis technique to predict clutch life with the help of numerous vehicle parameters such as transmission oil temperature, vehicle speed, vehicle torque, vehicle engine speed, transmission oil level, accelerometer pedal position, parking brake status, and oil contamination. The consideration of many parameters adds to enhancing the accuracy of the forecast output by expanding the number of parameters considered. The proposed system's performance, which has an accuracy of 94 percent, is considered satisfactory. This technology can be used to notify drivers in Lehman language about projected consequences depending on the information provided by the system.

1. INTRODUCTION

In response to stringent vehicle emission regulations, the automotive industry has encountered difficulties in the development of intelligent vehicle components, which contribute to the performance of vehicles and allow them to comply with pollution regulations. When determining emission regulations, a variety of factors may be taken into consideration. Engine operations must be carried out in a methodical manner in this context, as this will aid in the maintenance of the vehicle's emission system. Electronic systems play an increasingly essential part in the development of intelligent component systems for automobiles on a day-to-day basis. An active control system monitors the operation of a vehicle component by utilizing information provided by numerous sensors installed in the vehicle. On a global scale, vehicle emissions contributed significantly to the total amount of pollution produced. These emissions have a substantial impact on the ecosystem, and they are one of the most significant factors contributing to the environment's depletion of natural resources. Following as a result, a number of laws and pieces of legislation have been created to restrict emissions from motor vehicles. As a result of these rigorous emission requirements, intelligent design criteria for automobile components are established, which are advantageous in the regulation of vehicle emissions. The advancement in vehicle electronics systems, made possible through the use of electronic control unit (ECU) actuators, can collect data from a variety of smart components, including the anti-lock braking system (ABS), the engine management system (EMS), the instrument cluster (IC), the airbag, and the transmission system. With the availability of data from sensors from vehicle parts sensors, it is feasible to predict the life and status of parts in real time with the use of machine learning techniques using data from the sensors. It is possible that this type of prognosis study will contain statistical measures for anticipating the particular part status in terms of the performance attributes of that part.

Machine learning techniques for forecasting can be used to generate statistical analysis tools, which can then be applied to other data sets. This can be accomplished using previously recorded historical data. The primary goal of this work is to forecast clutch life with the help of a supervised dataset for training, which contains recorded information on a variety of different characteristics. Prediction accuracy is the primary performance measure examined during multiple regression analysis in the suggested method, and it is also the most important.

2. RELATED STUDY

The difficult task is to measure the status of car parts in the demanding domain of vehicle parts status measurement. These are some of the ways from among the many extant methods that have been developed for comparable prediction of part status. When it comes to predicting the clutch state, the paper [1] focuses on prognostic diagnosis through the application of

Bayesian algorithms and fuzzy rules. Paper [2] discusses machine learning approaches that can be used to make prognostic predictions, and it is available online. The prediction of the remaining useful life of sensors is conducted in article [3], as well. Prediction is accomplished by the application of deep learning techniques. On the other hand, paper [4] demonstrates proactive maintenance for transmission systems through the use of operations data and a neural network model. The applicability of regression models and classification models, as well as their corresponding performance evaluation metrics, are explored in paper [5]. It is the goal of this article to determine which machine learning method will be most effective in estimating the remaining useful life.

In paper [6], the forecast of battery life is accomplished through the application of regressive analysis. The method is used to anticipate the remaining useful life of lithium-ion batteries, and it is very accurate. In paper [7], a regression model is used to estimate the quality of changed oil, and the results are presented in Table 1. The values of interfacial tension are employed in the training and testing of the model, respectively. Different machine learning approaches are addressed in Paper [8], along with mathematical explanations and examples of their use in real-world situations. Book [9] provides theory of Artificial intelligence and machine learning.

Paper [10] illustrates fault diagnosis of vehicle steering actuator using the SVM approach. Paper [11] gives a fault detection approach for automated equipment using a machine learning algorithm. Paper [12] represent a predictive diagnostic system for a truck vehicle that helps to increase the reliability of transport and logistics system. Paper [13] illustrates acoustic-based engine fault detection. Paper [14] gives a case study approach for fault detection of an automated guided vehicle.

3. PROPOSED WORK

The system captures clutch system data via CAN-based protocols and sensors, and then stores it on remote servers using the Internet. Vehicle telematics will be used to collect and store data on the vehicle, which will be kept on a data server. This data must be preprocessed before it can be supplied into the machine learning computation component of the program. When the machine learning model is finished, the result will be sent to the display system.

In order to collect a variety of measured parameters through the network, the vehicle is equipped with a number of different sensors.

In a smart vehicle transmission system, many sensors are used to get system data. These sensors are engine oil temperatures sensors, vehicle speed sensors, and vehicle wheel speed sensors. Specification for engine oil temperature sensors such as OTS 001 and OTS 004, also specific vehicle wheel speed sensors such as 7G-Tronic.

The data gathering mechanism is depicted in Figure 1. An intelligent system can record the value of this sensor and use it later. The controller area network (CAN) is used to connect the ECU Intelligent system, which harvests data from all of the sensors, to the rest of the vehicle's electrical system. The data that has been extracted is saved so that it can be utilized for processing and forecasting.

In above figure of proposed system architecture basically 6 modules are considered. Named as vehicle sensor clutch and Electronics control unit (ECU), Gateway network, vehicle telematics, cloud based storage system, machine learning algorithm with multiple linear regression and ultimately prognostics user interface which gives output towards dashboard or tool display. Over vehicle sensor Electronics control unit (ECU) senses the clutch information which may be processed and delivered towards car telematics unit through controller area network (CAN). Further vehicle telematics transfer these data to the cloud based storage. This all stored data is preprocessed before machine learning model. After applying machine learning algorithm output is assessed via prognostic user interface. And ultimately this prognostics user interface which sends output towards car dashboard or vehicle OFF Board tool display.

In the proposed method collected data is split in 70%-30% parts. The 70% parts is used for training and 30% for testing. Data is split in 70-30% to get good overfitting of a predictive model which reduces mean error of training set and testing set.

The flow chart shown in Figure 2 explains the different steps covers to achieve machine learning operations for the predictions. Import all essential libraries like NumPy, matplotlib, and the specific algorithm library is the initial step in the system flow diagram shown above. Preprocessing and tweaking of the dataset follow as the next steps. Continue to read all datasets and generate training and testing data using the database as a basis. Fit method and input test data set prediction should be used first. Finally, compute the precision score and display a graphical representation of the dataset's confusion matrix.

Referring to the workflow diagram above the steps carried out as, First use all required libraries with set variables then preprocess the dataset with required filtering, missing data analytics, and define numeric status as required for machine learning operations. Next, divide and create testing and training data for the machine learning operations. Next, apply the algorithm on the dataset with the fit method of the SK-learn library. Then apply the fit method to training and testing data. Input test data can be set for prediction purposes. Measure result with confusion matrix and finally show the result with graphical charts.

Table 1. Clutch life status parameters

Parking Brake	Environment Temp °C	Vehicle Speed Km/h	Engine Torque Rpm	Transmission Oil Temp °C	Clutch Life
0	35	40	1500	49	1
0	39	25	1000	44	1
0	35	65	3000	105	2
1	30	35	2000	88	2
0	45	80	2000	67	2
0	50	80	2500	77	2
0	35	65	3000	105	2
1	30	35	2000	88	2
0	25	46	3500	121	3
0	55	42	4500	140	4

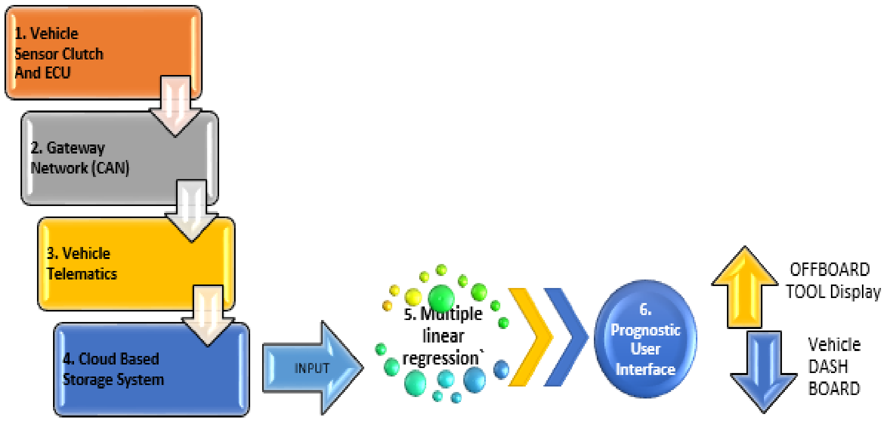

Figure 1. Proposed system architecture

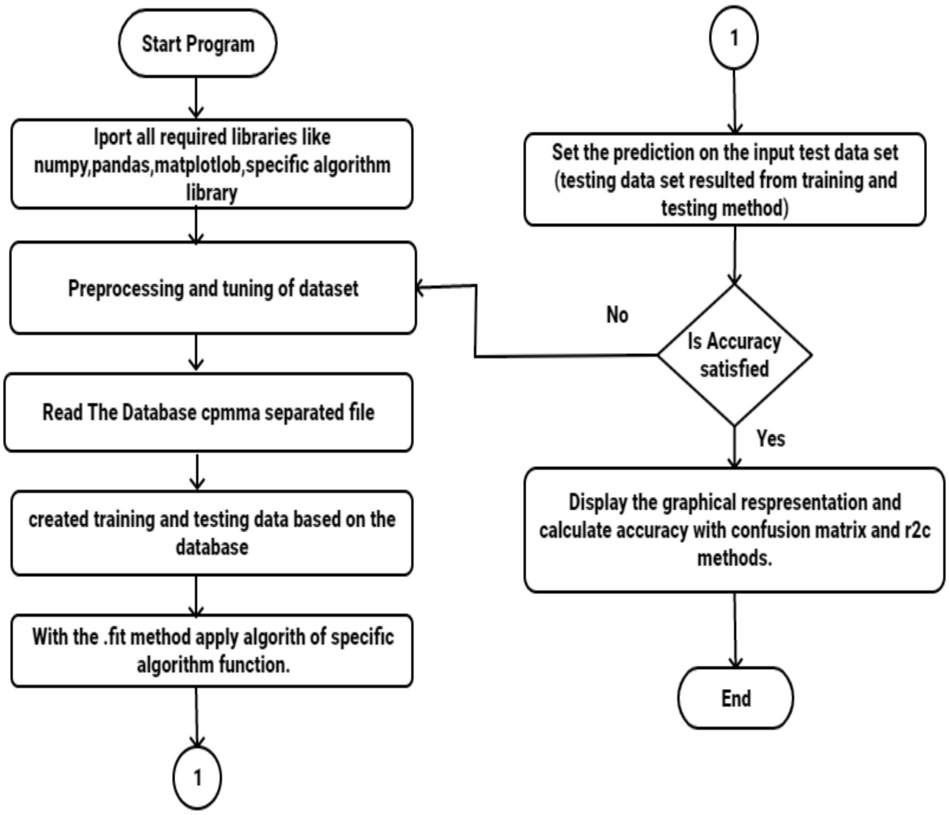

Figure 2. System flow chart

4. MULTIPLE LINEAR REGRESSION MODEL

Multiple linear regression is a subset of approaches to supervised machine learning that are utilised in a wide range of applications, including financial forecasting. The most common applications for which it is developed and utilised are classification, regression, and outlier identification. The multiple linear regression technique is utilised to solve the two-class classification problem that was presented during the experimentation phase.

In linear regression, only one parameter is considered whereas in practice other parameters are also responsible for accurate prediction. Multiple linear regression supports to consideration of multiple parameters which gives more meaningful and accurate predictions in line with vehicle operations.

Multivariable dependent regression is a type of statistical analysis in which every input X is coupled with every output Y, with a Y to X dependency between the two variables being studied. X1, X2, ..., Xp are the explanatory factors for the population regression line for P explanatory variables.

$$M_y = B_0 + B_1 x_1 + B_2 x_2 + \ldots + B_{p\ x} P \tag{1}$$

The hyperplane that distinguishes the variables has different responses for M variables than the other hyperplanes. The observed values for y fluctuate as a function of the variable My. For all possible values, the standard deviation is linear. In this case, the fitness of parameters based on deviation is sufficient, and the B0, B1, B2, ..., Bp parameters are estimated using population regression.

On the cross-section of the hyperplane crossed by X, Figure 3 shows a prediction of the n-dimensional data vector Y. This graphic depicts the column space of X that may be reached by

Y on the hyperplane by means of the X and Y-axis. Similar to the equation of the line, this graphic depicts the mathematical X-Y space covered by many regression lines in a similar manner.

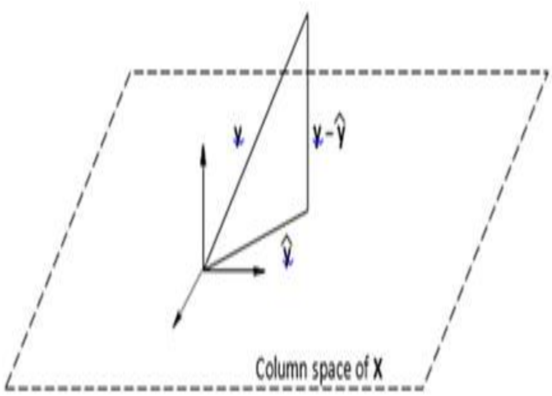

Figure 3. Multiple regression model

5. RESULT ANALYSIS

Different characteristics for the provided system are taken into consideration in the dataset, including the parking brake, the environment temperature in degree Celsius, the vehicle speed in kilometers per hour, the engine torque in revolutions per minute, the transmission oil temp in degree Celsius, and the clutch life. Crankshaft life, on the other hand, is a dependent parameter, whereas all other characteristics are independent. Table 1 shows the subset of the dataset that was fetched and displayed below.

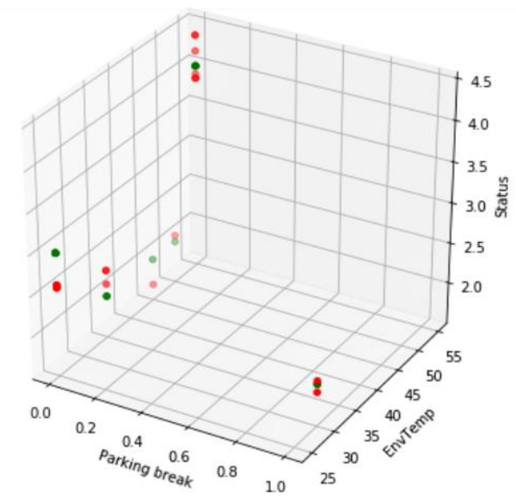

Figure 4. Graph 1. For clutch life status vs envtemp and parking brake

If you run a multivariate regression model on this dataset while taking into account two combinational factors, you will get three different regression lines in the output, as shown in the accompanying graphs for Graph 1, Graph 2, and Graph 3 (Figures 4-6). When it comes to clutch life status, Graph1 gives regression estimation of the parking brake and EnvTemp using the parking brake and EnvTemp variables. The regression estimation of vehicle speed and engine speed in relation to clutch life status is depicted in Graph 2 (below). For clutch life status, Graph 3 gives a regression estimation of transmission oil and vehicle speed based on the transmission oil and vehicle speed data.

Based on the examination of the three graphs above, it can be concluded that multiple regression for car clutch prognosis provides 94.5 percent accuracy for the system. Which lead to efficient model for predictions.

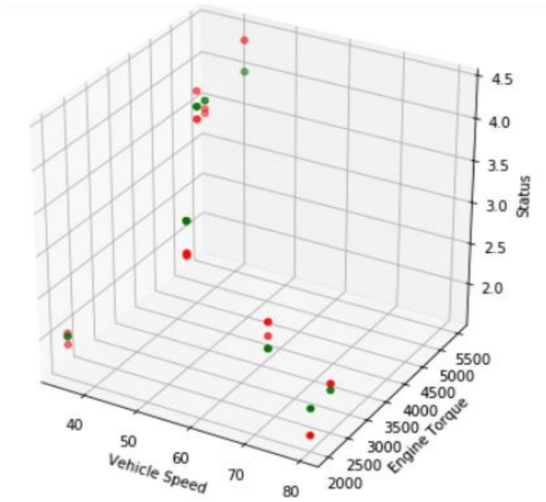

Figure 5. Graph 2. For clutch life status vs vehicle speed and engine torque

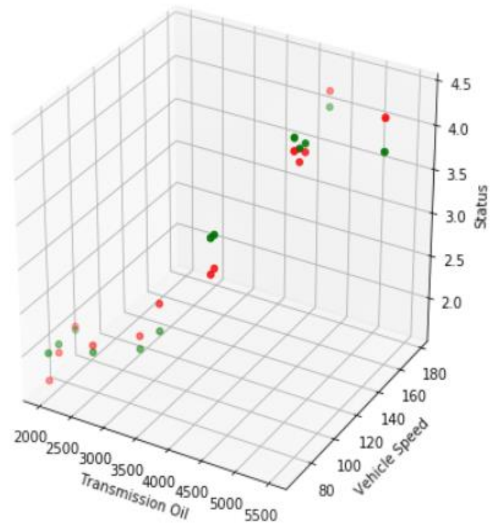

Figure 6. Graph 3. For clutch life status vs transmission oil and vehicle speed

Efficiency of system is also measured with recall score f1 score and precision score. The number of positive class predictions that are positive class predictions is measured by precision. The number of positive class predictions made from all positive examples in a dataset is measured by the recall. F-Measure generates a single score that combines precision and recall issues into a single value.

For the System of Breakdown Prediction of Vehicle Clutch Using Multiple Linear Regression, this algorithm gives Accuracy score as 94.5 Percent, Precision score as 87.9 percent, Recall score is 93.6 percent and F1_score is 90.3 percent. Following graph in Figure 7 depicts Accuracy, Precision, Recall and F1_score for Breakdown Prediction of

Vehicle Clutch.

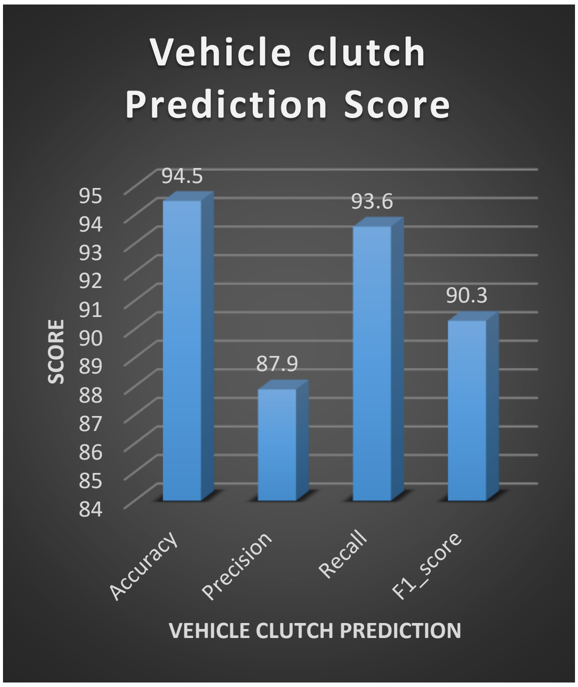

Figure 7. Vehicle clutch prediction score graph

6. FUTURE SCOPE

Machine Learning can extend the possibilities of various resolutions on available data through vehicle electronics. Predictive diagnosis will play an important and major role in the future. Vehicular operations and their defects as electronics involve major areas in the vehicle functionalities. This research work addresses the combinations of available vehicle data produced from the vehicle with a machine learning algorithm. In the future, such combinations of different algorithms can be designed to get the required vehicle prognostics operations.

7. CONCLUSIONS

For the purpose of predicting the status of the clutch, a multiple regression analysis was carried out. Various characteristics such as engine torque, parking brake, environmental temperature, engine speed, and transmission oil temperature are employed as independent variables, whereas the clutch status parameter is used as a dependent variable in this experiment.

Engine torque, parking brake, environmental temperature, engine speed, and transmission oil temperature are the important parameters that contribute to getting the vehicle clutch status thus these different parameters act as independent parameters whereas vehicle clutch status I dependent parameters by which clutch status has different meaningful states.

Using multiple regression analysis, the intercept and coefficient values are calculated, which allowed to anticipate the final output value. Considering a variety of scenarios, multiple regression analysis has been covered here. Using the present dataset sample 94 percent of accuracy is achieved. In order to yield useful information, raw data have been first preprocessed and mapped before being used. This data collection is used as an input to the machine learning model, which then learns from the data it has collected. Applied machine learning approaches in vehicle prognostics have a wide range of possible applications in a variety of fields. Developing deep learning algorithms for early flaw identification may be researched in the future, which would be advantageous for further research.

REFERENCES

[1] Vanjire, S.K., Patil, S.B. (2022). Research survey on machine learning used in vehicle prognostics. Journal of Analysis and Computation (JAC), 1-6.

[2] Vanjire, S.K., Patil, S., Rajankar, O. (2020). Predictive diagnostic approach for vehicle transmission clutch using fuzzy logic and Bayesian algorithm. TEST, 2102-2107.

[3] Wu, J., Hu, K., Cheng, Y., Zhu, H., Shao, X., Wang, Y. (2020). Data-driven remaining useful life prediction via multiple sensor signals and deep long short-term memory neural network. ISA Transactions, 97: 241-250. https://doi.org/10.1016/j.isatra.2019.07.004

[4] Furch, J., Krobot, Z. (2020). Trends in predictive and proactive maintenance of motor vehicles. Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis, 68(2): 311-322. http://dx.doi.org/10.11118/actaun202068020311

[5] Saranya, E., Sivakumar, P.B. (2020). Data-driven prognostics for run-to-failure data employing machine learning models. In 2020 International Conference on Inventive Computation Technologies (ICICT), pp. 528-533. https://doi.org/10.1109/ICICT48043.2020.9112411

[6] Botchkarev, A. (2018). Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology. arXiv preprint arXiv:1809.03006. https://doi.org/10.48550/arXiv.1809.03006

[7] Yin, S., Pang, J., Liu, D., Peng, Y. (2013). Remaining useful life prognostics for lithium-ion battery based on Gaussian processing regression combined with the empirical model. In Annual Conference of the PHM Society, 5(1). https://doi.org/10.36001/phmconf.2013.v5i1.2179

[8] Bhatia, N.K., El-Hag, A.H., Shaban, K.B. (2020). Machine learning-based regression and classification models for oil assessment of power transformers. In 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), pp. 400-403. https://doi.org/10.1109/ICIoT48696.2020.9089647

[9] Joshi, A.V. (2020). Machine Learning and Artificial Intelligence. Springer Cham. https://doi.org/10.1007/978-3-030-26622-6

[10] Shi, Q., Zhang, H. (2020). Fault diagnosis of an autonomous vehicle with an improved SVM algorithm subject to unbalanced datasets. IEEE Transactions on Industrial Electronics, 68(7): 6248-6256. https://doi.org/10.1109/TIE.2020.2994868

[11] Basangar, S., Tripathi, B.N. (2020). Literature review on fault detection of equipment using machine learning techniques. In 2020 International Conference on Computation, Automation and Knowledge Management (ICCAKM), pp. 62-67. https://doi.org/10.1109/ICCAKM46823.2020.9051543

[12] Demyanov, D.N., Simonova, L.A., Kapitonov, A.A. (2020). The work algorithm of the truck intelligent predictive diagnostics system. In 2020 International Russian Automation Conference (RusAutoCon), pp. 621-625. https://doi.org/10.1109/RusAutoCon49822.2020.9208065
[13] Mathew, S.K., Zhang, Y. (2020). Acoustic-based engine fault diagnosis using WPT, PCA and Bayesian optimization. Applied Sciences, 10(19): 6890. https://doi.org/10.3390/app10196890
[14] Witczak, M., Mrugalski, M., Pazera, M., Kukurowski, N. (2020). Fault diagnosis of an automated guided vehicle with torque and motion forces estimation: A case study. ISA Transactions, 104: 370-381. https://doi.org/10.1016/j.isatra.2020.05.012