

PAPER • OPEN ACCESS

## Analysis on current flow style for vehicle alternator fault prediction

To cite this article: M A Abd Halim *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **670** 012042

View the [article online](#) for updates and enhancements.

# Analysis on current flow style for vehicle alternator fault prediction

**M A Abd Halim, M T A Rahman\*, N Abdul Rahim, A Rahman, A F A Hamid and N A M Amin**

School of Mechatronic Engineering, Universiti Malaysia Perlis (UniMAP), Kampus Pauh Putra, 02600 Arau, Perlis, Malaysia

\*Corresponding author: tasyrif@unimap.edu.my

**Abstract.** Vehicle alternator is only seen as fragment piece in vehicle. This project will analyse the vehicle alternator current output flow style. A study on charging rate onto battery can be made based on this analyst. From this a prediction can be made on the vehicle alternator health and may prevent it from affecting other charging system component. Features extracted from the raw sample data are root mean square (RMS), waveform length (WL) and autoregressive (AR). These features will then go through normality test to find the sample is normally distributed or not. The normality test used in this experiment is Jarque-Bera (JB) test. After go through the normality test, it shows that need to continue with non-parametric test. Because JB test shows that p-value is less than 0.05 confidence level. Kruskal-Wallis is used as non-parametric data validation. In this test, the hypothesized value is the value is the median instead of the mean as in Analysis of Variance (ANOVA). The Kruskal-Wallis test evaluates for any significance difference in the population medians on a dependent variable across all levels of a factor. For classification K-Nearest Neighbour (KNN) is used to find the number of K to differentiate between classes. After that the K value is use in Holdout method for training and testing. Final result shows that the accuracy of this machine learning tools is 94%. This number is a good percentage to be able to called as vehicle alternator fault prediction.

## 1. Introduction

Data prediction plays a tremendous trend nowadays. For now, quite hype with Industrial Revolution 4.0 movement where we can have real-time monitoring data in a cloud[1]. With the help of Artificial Intelligence, it will help enhance the movement. Data prediction is a very useful added value in any system where from it may contribute to help in maintenance and research and development. Fault prediction on vehicle alternator will help user from go through such event such as car breakdown during travelling. In order to make prediction a study and analyse on alternator output pattern. Later on then can make prediction on fault state of vehicle alternator[2], [3].

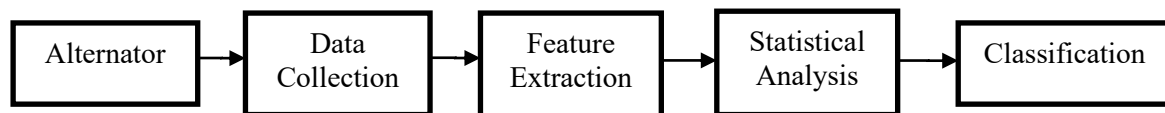
## 2. Methodology

This experiment is done on a Proton Wire 1.3 engine setup. On this experiment only load from the engine is test without others vehicle electrical load. A current sensor which is ACS 756 embedded with Arduino Uno microcontroller is used[4]. The sensor was connected to the terminal B on the alternator. This terminal B is where the generated output from the alternator[5]. The collected raw data as in Figure 1 was then computed into Matlab software to continue with statistical analysis and fault



Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

prediction. In statistical analysis, there are several tests that the raw data need to go through. First is Feature Extraction, here the feature that been extract is RMS, WL and AR. Then a normality test is done to find out the data is normally distributed or not. After that the feature data will go through data validation process which to find out which feature can be used for classification and which cannot[6]. Lastly classification and prediction, the data is then will be predict to find does it matching with the same class of it miss to other class.



**Figure 1.** Block Diagram.

### 2.1. Data Collection

The data was collected in several RPM which is on idle, 2000 and 3000. There are 3 type of alternator used, Alternator A is a good alternator, Alternator B which is a bad alternator suspected with rectifier problem and Alternator C which is suspected with IC regulator problem.

### 2.2. Feature Extraction

The feature extraction process is done in the Matlab. The raw data is being segmented to 100 sample data in 5s time data window which equivalent to 20hz frequency. The extracted features are RMS, WL, and AR. RMS, a meaningful way of calculating the average of values over a period of time. WL, the horizontal distance between successive crests, troughs or other parts of a wave. AR, a time series model that uses observations from previous time steps as input to a regression equation to predict the value at the next time step.

### 2.3. Statistical Analysis

Succeeding the features extraction process is statistical analysis. Statistical analysis consists of normality test and data validation. Normality test is use to find the acquired features is normally distributed or not. After normality test then can know what type of classification tool need to be used in data validation which is either parametric or non-parametric test. For this normality test, Jarque-Bera test is the chosen one among many other normality tests. It is simply because is a test for skewness and kurtosis and very effective. Jarque-Bera test or known as JB test is where a null hypothesis is tested with vector from the normal distribution with an anonymous mean and variance[7].

$$JB = \frac{n}{6} \left( s^2 + \frac{(k-3)^2}{4} \right) \quad (1.1)$$

Where n is the sample size, s is the skewness and k are the sample kurtosis. For larger sample size, the test statistics has a chi-square distribution with two degree of freedom. This normality test is to measure the p-value of all these features to find out where the significance level when set to 0.05. When p-value<0.05, null hypothesis can be rejected. The hypothesis for normality test are as follows:

$$\begin{aligned} H_0: & \text{Data is normal distribution} \\ H_1: & \text{Data is not normal distribution} \end{aligned}$$

From JB test, shows that it is accepted to reject the null hypothesis because the p-value<0.05 and shows the data is not normally distributed so need to continue the statistical analysis with non-

parametric test tools. All features are already analysed by Jarque-Bera test which is one of the time frequency domain analysis. For non-parametric data validation test, Kruskal-Wallis test is chosen. In this test, the hypothesized value is the value is the median instead of the mean (as in ANOVA). The Kruskal-Wallis test evaluates for any significance difference in the population medians on a dependent variable across all levels of a factor. If the requirements for a parametric test can be met, it is better to use the one-way independent measure ANOVA instead of Kruskal-Wallis since it is more powerful.

$$H = \frac{12}{N(N+1)} \left( \frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \dots + \frac{R_k^2}{n_k} \right) - 3(N+1) \quad (1.2)$$

Where  $n$  is the size of sample  $k$ ,  $N$  is the total number of observations in combined samples,  $k$  is the number of samples or groups and  $R_k$  is the sum of ranks from sample  $k$ .

$$H_0: \text{Median1} = \text{Median2} = \text{Median3}$$

$$H_1: \text{At least one group median is different from others}$$

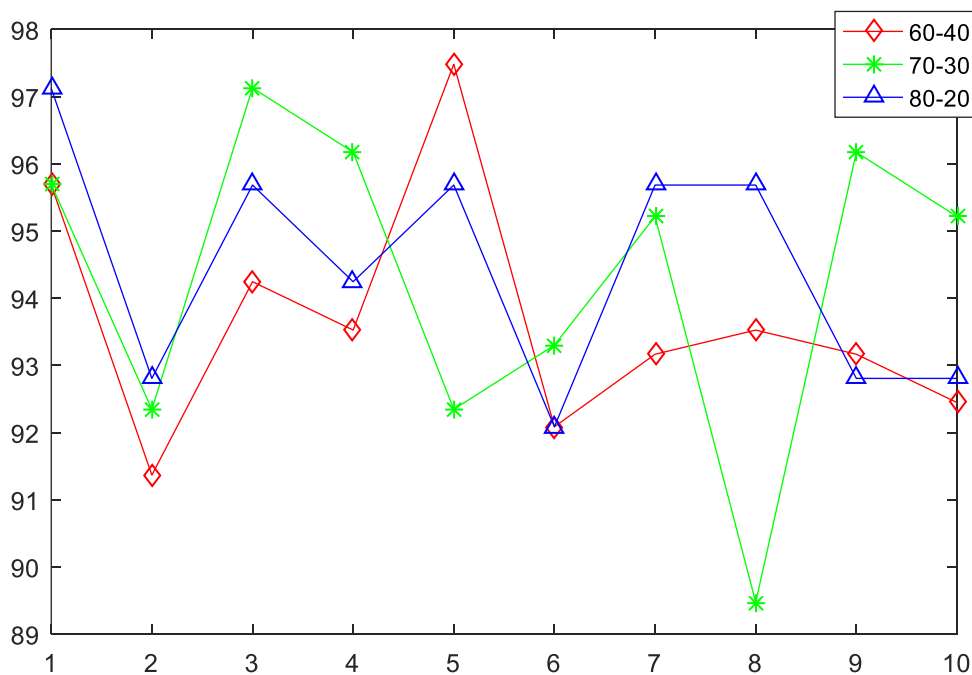
In Kruskal-Wallis test the to reject null hypothesis the p-value must less than 0.05. If it cannot reject null hypothesis means that the median for each group is the same and it may affect the classification accuracy later. Because it cannot differentiate what type of group is need to sort. For Kruskal-Wallis test, it shows that all features are rejecting null hypothesis except feature 6 which is AR4 feature where the  $p\text{-value} > 0.05$ . Here a conclusion may arise which all features the median is the same between classes but for feature 6 at least one group median is different from others. An early prediction can be made where feature 6 cannot differentiate each class and it may affect the accuracy. So, a new dataset form where feature 6 is remove from it and from 7 features it become only 6 features on it. For the next classification method both datasets is done where a finding on the affection of feature 6 for this project.

#### 2.4. Classification

In machine learning, KNN is used together with holdout method to do the prediction. KNN is a non-parametric, lazy learning algorithm. Its purpose is to use a database in which the data points are separated into several classes to predict the classification of a new sample point. KNN makes predictions just-in-time by calculating the similarity between an input sample and each training instance. The holdout method is the simplest kind of cross validation. The data set is separated into two sets, called the training set and the testing set. The function approximate fits a function using the training set only. Then the function approximate is asked to predict the output values for the data in the testing set[8]. The errors it makes are accumulated as before to give the mean absolute test set error, which is used to evaluate the model. The advantage of this method is that it is usually preferable to the residual method and takes no longer to compute. For this experiment, the holdout is used in 3 stages which is 60-40, 70-30 and 80-20.

### 3. Results and Discussions

In KNN the parameter to find classification is K value tested from  $K=1$  to  $K=10$ . For holdout method, it is tested for 10 times and the average accuracy is then accumulated in Table 1.



**Figure 2.** K-Nearest Neighbour

Figure.2 is the K value for KNN where it has been test from K=1 to K=10 and find the best accuracy. Later then holdout method will come around to find out the distribution accuracy. The holdout method is done for 10 times then get the average accuracy.

**Table 1.** Holdout Method.

Type	Holdout 60-40	Holdout 70-30	Holdout 80-20
7 Features	94.3885	94.4019	94.8921
6 Features	94.3525	94.1627	95.0360

Based on Table 1 result we may assume the result is still approximately same. This step may lead to new approach with is reduce complexity where feature 6 is remove from the analysis but still lead to equivalent result. From the result, it shows that the accuracy for this model to do the prediction on the vehicle alternator fault is around 94%. Even though the feature 6 is removed it still resulting an equivalent result and this may help in reduced complexity.

#### 4. Conclusions

Based on the normality test, the findings are that the data is not normally distributed. Jarque-Bera test is the chosen normality test. Therefore, it must go to non-parametric test. There are a lot of type of parametric test but the chosen one Kruskal-Wallis for the non-parametric data validation. The result from Kruskal-Wallis show that feature 6 does not reject the null hypothesis hence the median value for each class is approximately same. The sample data is then divided into 2 which one for all features and the other one with feature 6 exclude from it. This is because later in the experiment want to find did feature 6 affecting the accuracy. For non-parametric classification, K-Nearest Neighbour is chosen among other method. This is to find the number of K so the class can be differentiating. The test is done for 10 times and the number of K is trained by using cross validation technique which is Holdout method. For Holdout method there is 3 type of Holdout setting used. First with 60-40 holdout which

means 60% training and 40% testing. Next with 70-30 holdout which means 70% training and 30% testing. Last but not least 80-20 holdout which means 80% training and 20% testing. All holdout setup was being run for 10 times and the average value is recorded. As a conclusion, the result from this method find that the accuracy for all 7 features intact were 94% and for only 6 features were also 94%. From this result can conclude that feature 6 which is AR4 does not affect the classification accuracy. So an addition conclusion that can be made that feature AR4 can be removed from the features extracted and this may result reduced complexity because it does not affect the accuracy much.

### Acknowledgement

The authors would like to extend their sincere gratitude for the funding provided by the Universiti Malaysia Perlis (UniMAP) (Geran Penyelidikan Khas - No.9004-00064) which enable this research to be carried out.

### References

- [1] Hamid A F A et al. 2017 Connected car: Engines diagnostic via Internet of Things (IoT) *J. Phys. Conf. Ser.* **908** (1).
- [2] Şimşir M, Bayir R and Uyaroğlu Y 2016 Real-time monitoring and fault diagnosis of a low power hub motor using feedforward neural network *Comput. Intell. Neurosci.* **2016**.
- [3] Bangalore P and Tjernberg L B 2015 An Artificial Neural Network Approach for Early Fault Detection of Gearbox Bearings *IEEE Trans. Smart Grid* **6** (2) 980–987.
- [4] Dinio C J T et al. 2019 Automated water source scheduling system with flow control system *2018 IEEE 10th Int. Conf. Humanoid, Nanotechnology, Inf. Technol. Commun. Control. Environ. Manag. HNICEM 2018*.
- [5] Mazlan R, Dan R and Zakaria M 2017 Experimental study on the effect of alternator speed to the car charging system *The 2nd International Conference on Automotive Innovation and Green Vehicle (AiGEV 2016)*.
- [6] Sukor A S A, Zakaria A and Rahim N A 2018 Activity Recognition using Accelerometer Sensor and Machine Learning Classifiers *2018 IEEE 14th Int. Colloq. Signal Process. Its Appl.* 233–238.
- [7] Reddy D N and Ravinder Y 2018 Novel receiver diversity combining methods for spectrum sensing using meta-analytic approach based on p-values *J. Eng. Sc. Tech* **13** (9) 2883–2897.
- [8] Yadav S 2016 Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification *2016 IEEE 6th Int. Conf. Adv. Comput.* 78–83.