Machine Learning Based Vehicle Performance Analyzer

ABSTRACT:

Predicting the performance level of cars is an important and interesting problem. The main goal is to predict the performance of the car to improve certain behaviours of the vehicle. This can significantly help to improve the system's fuel consumption and increase efficiency. The performance analysis of the car is based on the engine type, no of engine cylinders, fuel type, horsepower, etc. These are the factors on which the health of the car can be predicted. It is an ongoing process of obtaining, researching, analyzing, and recording health based on the above three factors. The performance objectives like mileage, dependability, flexibility and cost can be grouped together to play a vital role in the prediction engine and engine management system. This approach is a very important step towards understanding the vehicle's performance.

INTRODUCTION:

After more than one hundred years of development, the automobile has become an indispensable means of transport in human life. Automobiles have also become the pillar industries in social and economic development. In the increasingly sophisticated automobile consumption, the performance of automobile has caused widespread concern. Vehicle performance (overall performance) are related to a number of factors such as economy, power, handling stability, environmental protection, safety, comfort and others. Due to the differences between the correlation of the various factors on the performance of automobile are difficult to determine, apparently according to the evaluation of individual indicators may result in poor evaluation of the compatibility issues Therefore, it is necessary to establish a scientific, reasonable and simple method for evaluating the performance of the entire vehicle. The evaluation of the target vehicle performance by using the comprehensive performance evaluation model based on set theory can overcome the problems of many factors and different degrees of influence affecting the vehicle performance.

PROBLEM STAEMENT:

The vehicle industry is extensively using manually engineered knowledge based methods for diagnostics and prognostics. These are development resource intense and limits the adoption to systems which are expensive, safety critical or under legal obligation of monitoring. The demand for more uptime requires less unplanned maintenance which in return drives maintenance predictions. This requires more universal deployment of diagnostics or maintenance predictions as the systems that today are under diagnostic monitoring only account for a small fraction of all repairs.

PROBLEM STATEMENT SOLUTION:

We will be able to understand the problem to classify if it is a regression or a classification kind of problem and we can able to know how to pre-process/clean the data using different data pre -processing techniques. People will be able to analyse or get insights into data through visualization. Applying different algorithms according to the dataset and based on visualization. Finally will be able to know how to build a web application using the Flask framework.

FEATURES:

- Fuel economy & emissions
- Thermal and energy management
- NVH and acoustics, pass-by-noise regulation
- Durability
- Drivability
- Driving dynamics
- Integrated safety
- Aerodynamic performance
- Water management

METHODOLGY PROPOSED:

The collection of dataset was manually done from various automobile websites. The missing values (Noise) was removed and the clean data was used for selection of attributes. PC analysis was done and regression technique was applied using data mining tool.

EXISTING METHOD:

Methods used for vehicle detection can be divided into two groups - intrusive and non-intrusive. Intrusive sensors include inductive loops, magnetometers, micro-loop probes, pneumatic road tubes, piezoelectric cables and other weigh-in-motion sensors. These devices are installed directly on the pavement surface, in saw-cuts, in road surface holes by tunneling under the surface or by anchoring directly to the pavement surface as vehicle detection is an important process in intelligent transport management systems, as it allows the collection of big data on vehicles' speeds and weights as well as the traffic intensity, which helps to enhance smooth transportation and to reduce road accidents. Non-intrusive and intrusive sensor technologies are often employed for monitoring. A laser sensor, a temperature sensor, or an image-based sensor have been used for vehicle detection that is based on a change in laser light intensity, temperature value, or imaging property due to the vehicle appearance.

IMPLEMENTATION TECHNIQUE DETAILS:

1.Data Collection

ML depends heavily on data, without data, it is impossible for an “AI” to learn. It is the most crucial aspect that makes algorithm training possible. In Machine Learning projects, we need a training data set. It is the actual data set used to train the model for performing various actions.

2. Data Pre-processing

Data pre-processing is a process of cleaning the raw data i.e. the data is collected in the real world and is converted to a clean data set. In other words, whenever the data is gathered from different sources it is collected in a raw format and this data isn't feasible for the analysis.
Therefore, certain steps are executed to convert the data into a small clean data set, this part of the process is called as data pre-processing Follow the following steps to process your Data

- Import the Libraries
- Importing the dataset
- Taking care of Missing Data
- Label encoding
- One Hot Encoding
- Feature Scaling
- Splitting Data into Train and Test

3. Model Building

There are several Machine learning algorithms to be used depending on the data you are going to process such as images, sound, text, and numerical values. The algorithms that you can choose according to the objective that you might have it may be Classification algorithms or Regression algorithms.

4.Application Building

In this section, we will be building a web application that is integrated into the model we built. A UI is provided for the users where he has to enter the values for predictions. The enter values are given to the saved model and prediction is showcased on the UI. Previously we are saved this file as "regression.pkl". We have seven independent variables and one dependent variable for this model.

RESULT AND DISCUSSION:

The speed and distance of a vehicle can be monitored easily with the help of speed sensor. There are many types of speed sensors available. Proximity sensor is the best one to measure the speed of the car. The proximity sensor will work with magnetic pick up , the metal piece of the tyre drum should cut the flux in between the wheel & sensor for each rotation. For one rotation one pulse will be taken and then the output will be given to a microcontroller, and the microcontroller output will be calculated for a minute & hour we can calculate the KMPH & the distance. For this we need a timer programming concept in the microcontroller. We performed several experiments on the parameters which are essential for the testing of vehicle's safety and economic efficiency. In our first experiment, a relationship between Maximum speed value and the travel time (red zone) is obtained. This relationship describes the total distances travelled while crossing the road given the signal was red

IMPROVEMENTS:

This project is designed with Microcontroller, LCD display, Temperature sensor, Pressure sensor, Fuel sensor, Speed sensor, SCU, Matrix keypad, Amplifier and ADC. The whole setup is to be placed on the dashboard of the vehicle. In this block the speed of the wheel, the battery failure indication is done along with which the fuel level is also included, temperature of engine then pressures of oil. We are able to calculate the distance covered by the vehicle. Moreover, the speed of the vehicle is calculated in kilo meter /hour by using some mathematical calculation.

TABLES:

TABLE 1: Correlation matrix of SFTY_DRVG_INDX

SFTY_DRVG_INDX	Correlation Matrix	
	Features	**Correlation Value**
ROF_XCH_SCR	Fuel Efficiency	1.000000
AVG_SPD_VAL	Average Speed Value	0.951148
MAX_SPD_VAL	Maximum Speed Value	0.850964
SPD_TH4_DRVG_TIM_VAL	Fourth Throttle Driving Time Value	0.676948
IDL_HCT	Idle Time Value	-0.296508

Table 1 shows the correlation values for our hypothesis #1, Safe driving Index. And Table 2 shows the correlation values we used for our hypothesis #2, Economic driving Index (ECN_DRVG_INDX). We found that Fuel Efficiency in the following table showed a 100% correlation with our hypothesis #1, but considering it alone for the training would lead to over fitting of the data and eventually reduce the efficiency of the model. Therefore, it was important to consider all the features.

TABLE 2: Correlation matrix of ECN_DRVG_INDX

ECNM_DRVG_INDX	Correlation Matrix	
	Features	**Correlation Value**
TH5_RPM_TIM	Fifth Throttle RPM time	-0.567331
UGY_ACSD_OFT	Urgent Acceleration Number	-0.615989
UGY_RDSD_OFT	Urgent Deceleration Number	-0.621209
TH4_RPM_TIM	Fourth Throttle RPM time	-0.563859

CHARTS:

We performed a k-fold cross-validation technique [6] with k=10, to train the model. In the Hypothesis-1, ECN_DRVG_INDX, we found that majority of the data was congested in the lower left part of the graph suggesting an inverse logarithmic growth of the trend based on the training data. This showed a positive growth of the ECN_DRVG_INDX based on the hypothesis value. However, the data scatters as the value on x-axis increases, hinting at a somewhat lesser correlation for predicted value based on hypothesis value. Therefore, the ECN_DRVG_INDX is found to be an inverse logarithmic function of the features.

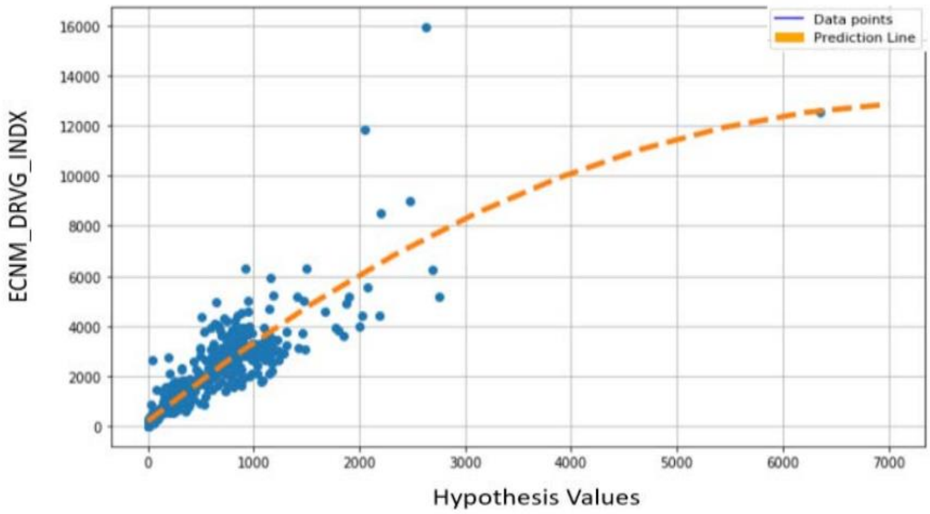
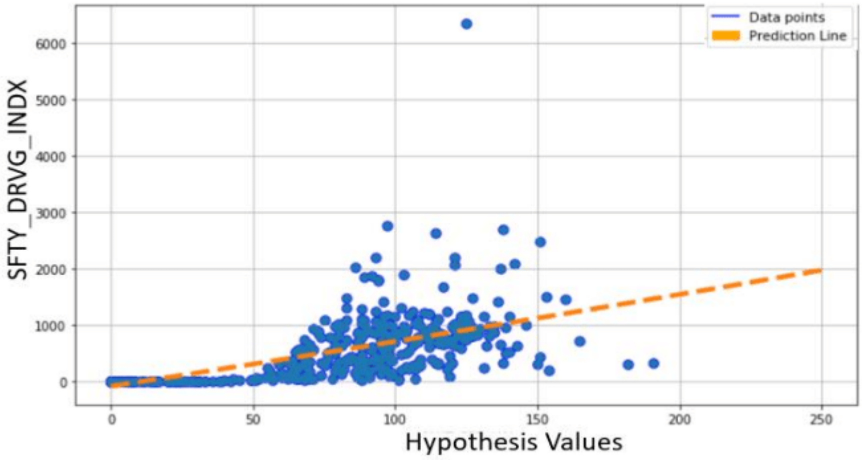


Figure 1: Linear regression based prediction trend of ECN_DRVG_INDX against hypothesis value.



In the Hypothesis-2, safety driving index (SFTY_DRVG_INDX), we observed a slower initial growth in the trend but rapidly picks up the momentum after crossing a certain threshold value. After this value the trend moves in a logarithmic fashion up towards the positive y-axis. We do observe few outliers that were safely ignored for consideration. Therefore, the safety driving index (SFTY_DRVG_INDX) is found to be a logarithmic function of the features considered in our hypothesis. However, a good example for future work would be to obtain different datasets on vehicle metrics and train our model on that data to further improve the accuracy from 80% to more than 90%. However, an accuracy of 80% in the economic driving index, which represents a car's health condition, is good enough to include a human-in-the-loop for the prediction work. We aim to implement our model in an IOT system for driving school scenario to judge the driver driving ability and frame a standard whether the driver should be eligible to obtain a driver's license or not.

INNOVATION:

This project is designed with Microcontroller, LCD display, Temperature sensor, Pressure sensor, Fuel sensor, Speed sensor, SCU, Matrix keypad, Amplifier and ADC. The whole setup is to be placed on the dashboard of the vehicle. In this block the speed of the wheel, the battery failure indication is done along with which the fuel level is also included, temperature of engine then pressures of oil. We are able to calculate the distance covered by the vehicle. Moreover, the speed of the vehicle is calculated in kilometer /hour by using some mathematical calculation.

DESCRIPTION ABOUT NEW FINDINGS

This project is designed with Microcontroller, LCD display, Temperature sensor, Pressure sensor, Fuel sensor, Speed sensor, SCU, Matrix keypad, Amplifier and ADC. The whole setup is to be placed on the dashboard of the vehicle. In this block the speed of the wheel, the battery failure indication is done along with which the fuel level is also included, temperature of engine then pressures of oil. We are able to calculate the distance covered by the vehicle.Moreover, the speed of the vehicle is calculated in kilometer /hour by using some mathematical calculation

CONCLUSION:

The results have proven to be approximately 80% fitting the given features and are very helpful to be used in different use cases such as a parameter in finding the driver's driving performance in a driving school, as a good estimate for finding an optimal price for a used car that can be based on several factors which we have analyzed in this paper etc. We also found that the model used to train the data can be improved further by finding better hyper parameter values for the features. It is also possible that different features can be considered for improving the hypothesis.

REFERENCES:

[1] Singh D, Singh M., "Internet of Vehicles for Smart and Safe Driving", International Conference on Connected Vehicles and Expo (ICCVE), Shenzhen, 19-23 Oct., 2015.

[2] Zhang, Y., Lin, W., and Chin, Y., "Data-Driven Driving Skill Characterization: Algorithm Comparison and Decision Fusion," SAE Technical Paper 2009-01- 1286, 2009, https://doi.org/10.4271/2009-01-1286.Azevedo, C. L Cardoso.

[3] J. E. Meseguer, C. T. Calafate, J. C. Cano and P. Manzoni, "DrivingStyles: A smartphone application to assess driver behavior," 2013 IEEE Symposium on Computers and Communications (ISCC), Split, 2013, pp.000535-000540. doi: 10.1109/ISCC.2013.6755001.

[4] Schneider, A., Hommel, G., & Blettner, M. (2010). Linear Regression Analysis:Part 14 of a Series on Evaluation of Scientific Publications. Deutsches ÄrzteblattInternational, 107(44), pp. 776–782.