Cloud based Vehicle Breakdown Prediction and Monitoring Driver Behaviour

Kapil Sharma[1] Chinmay Tompe[2] Pavitra Vichare[3] Aradhana Manekar[4]
[4]Assistant Professor
[1,2,3,4]Department of Electronics & Telecommunication Engineering
[1,2,3,4]Thakur College of Engineering & Technology, Mumbai, India

Abstract— The usage of vehicles all over the worlds has drastically increased during the last decade. Over 60 million passenger cars have been manufactured in the year of 2012. This rapid increase of vehicles has led to many concerns for a range of people and organizations where most of these concerns are common to all parties. For example, all parties (i.e., drivers, insurance companies, fleet vehicle managements, low enforcements authorities, etc.) are concerned about reckless driving and driver anomalies whereas drivers as individuals and people who are willing to purchase and sell cars are concerned also about the condition of the vehicle. Given the potential benefits of vehicular data analysis and the availability of technologies such as OBD, several vehicle monitoring and intelligent transport systems have been proposed. But, in almost all the systems proposed, there has been either simple or no processing of the data gathered from the Engine Control Unit (ECU) prior to displaying and are restricted to monitoring. Also they, hardly include a backend; therefore, are limited by the computational power of the smart phone. Hence, it is hard to predict any undesired outcome, such as an accident or a failure of sensor since they require real time and long-term analysis of data regarding the driving habits and the vehicle condition. In the proposed system i.e. Cloud Based Driver Monitoring and Vehicle Breakdown Prediction using data analytics, the task starts with collection of data from the cloud bifurcation and categorization of data for ease of use. The categorized data is to be passed through a machine learning process. In Machine Learning we plan on using linear regression algorithm for training the model. Amazon Web Services will be used for fulfilling the requirements of database. Bifurcated and categorized data will be available to the user on an Android application which the user will have on his/her mobile phone. Using this application user will be well updated about various parameters of his vehicle whenever is car is being driven. The analyses are performed both in real time as well throughout a long period of time. While some of these analyses are performed within the app, more complex and resource consuming ones are performed in the back end.

Keywords: ECU, Vehicle Breakdown, Cloud

I. INTRODUCTION

Before zeroing on the importance it is beneficial to possess a basic knowledge of data bases and machine learning. With beforehand information the terms and clauses in the implementation process can be grasped in a better way with more amount of depth.

Usually the data which is obtained from sources are in binary and other symbolic form. This is not possible to understand for the layman user, hence analysis and categorization of data becomes an important factor in vehicle sensor analysis, for doing this we need to shift the data obtained from the vehicle to a database cloud from where it can be picked up by the respective people for further use and making it go through the machine learning process.

The task starts with using Amazon Web Services a data base where the data is sent in binary or symbolic form. This data is now collected from AWS and categorized into a language known by the user either by generalizing the symbols or using any other categorization algorithm. There on we use machine learning, linear regression algorithm specifically for analysis. Further this data is interfaced with an android application where the updated data can be seen and known by the user. The android application interfaced with the system is built by using Android Studio software. Various data like engine rpm, maf sensor values and acceleration can be seen on the user application.

This whole system can be used by owners of cars for staying updated regarding the vehicle system and will in turn prove to be useful for keeping maintenance records. If we go on for large scale applications this can be used by cab aggregators for keeping a record and watch on all their cars running in different places, it will make their task very easy and will provide a helping hand in giving company ratings to drivers and maintaining records regarding performance of all cars.

II. METHODOLOGY

We collected the data for vehicle from OBD-II port and to train the data we downloaded csv file from Kaggle site. It includes various parameters for multiple vehicle taken at different time. We plot the MAF and Engine RPM to see the relationship between them which is linear. Next step is to calculate gradient and then store it along with timestamp to use it for Linear Regression training input.

This gradient and timestamp is used in the Linear Regression algorithm to calculate the predicted date along with the graph. (It is shown in the result section)Timestamp is measured in the format of UNIX time format. UNIX time format is the number of milliseconds that has elapsed since the day 1st January 1970.To convert it to human date we use time library from python function used is time.ctime().This date is then stored in the date.txt file which is uploaded to the S3 bucket (result section).Driver behaviour is calculated from acceleration streams taken from the OBD dataset and the difference is calculated, if the difference is greater than 0.01 m/s^2 then we will consider it Reckless. Also when the speed difference is greater than 10 m/s we consider Reckless or else Normal. x_value and y_value are the acceleration in the direction of x-axis and y-axis respectively. A sudden change in the acceleration in any direction is a sign of Reckless Driving. Driver behaviour is also written in the driver.txt file which is also uploaded on the S3 bucket (shown in the result section).To use S3 service we need to create a user in IAM

role which will generate Access Key and Secret Key. We need to provide S3 permissions to that user and then save it. Then at S3 service we need to create a bucket and create a folder where all of the data will be uploaded from python and retrieved by the android application. (shown in the Result Section)In Android Studio we generated three activity viz. Main Activity, Vehicle Breakdown and Driver Monitoring. First Activity which is Main Activity is to display two buttons for two options for the user to choose. Second Activity is of Driver Monitoring, when we click on the button the application moves from Main Activity to Driver Monitoring Activity using Intents. Here, we need to provide the S3 access key and secret key to access the bucket where the driver.txt file is saved. Buffered Reader class is used to read the driver.txt file and display it on the Reason Section. The word Reckless or Normal is highlighted accordingly which is shown on the screen. Third Activity is used to display the graph that is being plot and stored in S3 bucket. The image is downloaded from the S3 bucket and is displayed on the screen. The date that is stored by the python program in date.txt is read into the app and is displayed in the Text View in the predicted date section. To run the app we need to use USB cable to install the .apk into our mobile phone All the services are connected to the android application through Mobile Hub services which gives each user an identity through Cognito Pool Identity Service.

III. PROPOSED WORK

All the data is from previous vehicle trips is being stored in a csv file that is available online from Kaggle site[1]. The data consists of parameters of different vehicle taken at different time. Parameters includes timestamp, engine rpm, MAF sensor data, throttle position, etc. Some of the data includes null value because of some error margins created in the data collection. The data is collected from the OBD-II port and is stored in the csv file. This file is used by our project to calculate the vehicle breakdown prediction and identify the driver as reckless or not. We are designing an android application that will display the date and the graph predicting the date of the MAF sensor with respect to time. This data is being calculated using Linear Regression algorithm in machine learning Linear Regression is a Machine Learning Algorithm that used for predicting a value based previous supplied data. It is a supervised learning algorithm as it requires labelled data and testing is done on the previous data. It uses a simple slope intercept formula y=m1x+c1, where m1 is the slope and c1 is the y-intercept [2]. The independent variable is x and the dependent variable is y, so the input will be x and the output will be y. Linear regression takes between two linearly related data. Examples are such as predicting no. of students passed on studying a particular hours or predicting weather. First the data is being collected in a csv file with all the parameters of the vehicle. We select Engine RPM and MAF value and plot a graph, we will see a linear relationship between them. This indicates that both the parameters are directly proportional with each other. After that we will take the gradient of Engine RPM and MAF value and store it along with the timestamp in a csv file using python program. Then we will plot the graph of the gradient and time. Linear regression is applied on this data as we will get a graph from matplotlib (Python module). This linear line is then used to calculate the sensor breakdown when the graph reaches zero. The value of x is taken for the value of y=0, that gives us the predicted date in timestamp. The timestamp is then stored in the notepad and then uploaded to AWS service S3 [3]. Android Application retrieves this data into the app and displays it on the Application.

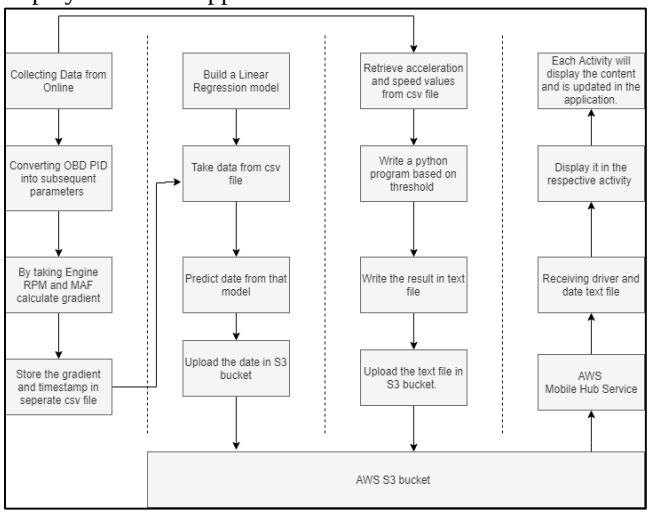

Fig. 1: Block Diagram of Project

IV. RESULTS

A. *MAF RPM gradient and Timestamp graph 2*

This is the graph that is plotted against timestamp (x -axis) and Gradient (y-axis) indicates a decreasing curve which means that the sensor is failing with time but it is taking longer time to reach zero. This data is taken from car 2 and the time is in Unix time format.

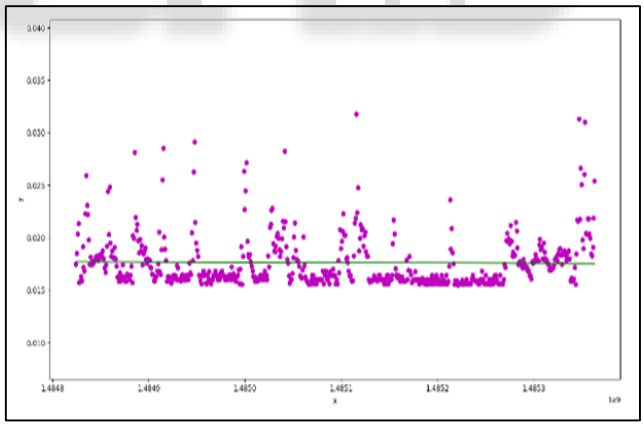

Fig. 2: MAF RPM gradient vs timestamp 2

B. *Android Application (Driver Behaviour)*

The behaviour is highlighted based upon the result we generated from python and the android application takes it from S3[4]. The reason for the behaviour is printed.

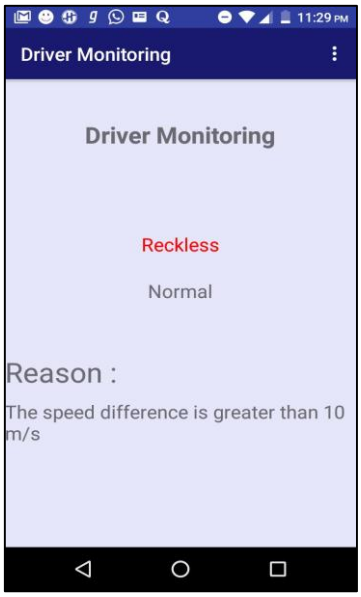

Fig. 3: Driver Monitoring Activity

C. Android Application (Vehicle Breakdown)

The output shows the graph that was plotted between the MAF RPM gradient and timestamp. It tells us when the sensor will be out of service (when the line reaches zero).The predicted date converted and stored by python program at S3 is displayed below [5].

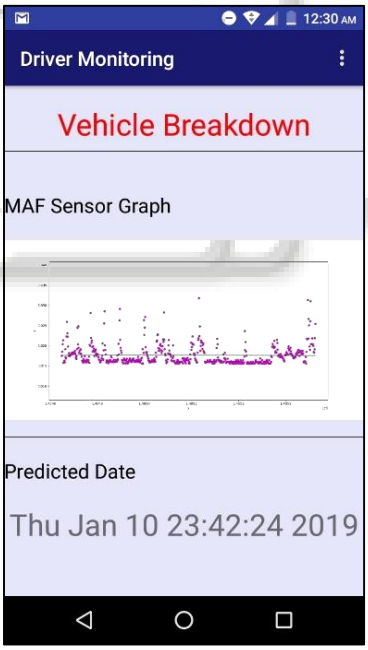

Fig. 4: Vehicle Breakdown Activity

V. CONCLUSION

In this project, we will take data from the OBD II port of the car and perform Data Analysis on the vehicle data that will be stored in the MYSQL relational database. The data will be bifurcated into proper column according to their type. This data can be used to train the machine learning model that will give us the habits of driver whether the driving is safe or not. We will also find the relation between fuel consumption and speed of the vehicle using machine learning algorithm Linear Regression.

Taking into consideration, the drawbacks stated in literature survey as well as in problem definition we have proposed a system wherein we provide analyzed and structured data. The diagnosis program displays the real-time conditions and the useful sensor data on the navigation screen. It also helps the driver to be aware of the vehicle conditions. A database management system is implemented for the storage and management of transmitted data and a graphical user interface (GUI) is developed for analyzing the transmitted data.

In the proposed system, Cloud is used as a remote server to receive all the transmitted data through wireless communication over internet connection. The data acquired is in the CSV file format. Microsoft excel is the most common program that is used to open as well as edit a CSV file. But these data received is meaningless and not user-friendly and needs to be decoded. The decoding of the data is done by Python program. It gives a structured format to these data. It imports CSV files and converts it into table or any other file format.

This pre-processed data is used to train the machine learning ng model that uses linear regression algorithm that predicts driver habits and relation between various parameters. This predicted result is displayed on android application.

VI. REFERENCES

[1] J. Varia and S. Mathew, "Overview of amazon web services," Http//Docs. Aws. Amaz. Com/Gettingstarted/Latest/Awsgsg-Intro/Intro. Html, no. December, 2013.
[2] K. H. Zou, K. Tuncali, and S. G. Silverman, "Statistical Concepts Series: Correlation and Simple Linear Regression," Radiology, vol. 227, pp. 617–622, 2003.
[3] M. Amarasinghe, S. Kottegoda, A. L. Arachchi, S. Muramudalige, H. M. N. Dilum Bandara, and A. Azeez, "Cloud-based driver monitoring and vehicle diagnostic with OBD2 telematics," IEEE Int. Conf. Electro Inf. Technol., vol. 2015-June, pp. 505–510, 2015.
[4] Wahab et al., "Driver Recognition System Using FNN and Statistical Methods," in Advances for In-Vehicle and Mobile Systems: Challenges for International Standards,Springer, 2007, ch. 2, sec. X, pp.011-023.
[5] S. Jain and M. A. Alam, "Comparative Study of Traditional Database and Cloud Computing Database," Int. J. Adv. Res. Comput. Sci., vol. 8, no. 2, pp. 80–87, 2017.

