Advanced Research in Energy and Engineering, Vol. 2 No. 1 (2020) p. 1-9

AREE

FAZ Publishing

Advanced Research in Energy and Engineering

Journal homepage: www.fazpublishing.com/aree

e-ISSN : 2682-8197

Vehicle Speed Prediction based on Road Status using Machine Learning

Mosabbir Bhuiyan[1,a], Md. Ahasan Kabir[2,b*]

Department of Electronics and Telecommunication Engineering[1,2]
Chittagong University of Engineering and Technology[1,2]
Chittagong, Bangladesh.

*Corresponding Author

Email: [a]mosabbirbhuiyanonady@gmail.com, [b]kabir.ece07@gmail.com

Received 10 January 2020;
Accepted 25 February 2020;
Available online 20 March 2020

Abstract: The prediction of vehicle speed plays a significant role in energy management. Predicting vehicle speed can help to designs a wide range of vehicle controllers, especially in case of an automated self-driving car for efficient fuel management applications. Nowadays the issue of energy saving is becoming more popular. This paper has proposed the prediction of the speed of the vehicle in different road statuses like in road curvature, in traffic, in weather, and indifferent road conditions, etc. Data is collected from different sources. This research is based on a supervised learning algorithm. Under supervised learning, a linear regression algorithm has been used to train up the model. For implementation Python programming language has been used. The proposed algorithm has provided better accuracy and performance than the existing well knew state-of-arts algorithm.

Keywords: Fuel Consumption, Linear regression algorithm, Machine Learning, Road condition, Vehicle Speed Prediction.

1. Introduction

The recent increase in fuel price is having a big impact on global economic developments. Everyone is worried about fuel consumption especially vehicle drivers. Not only does heavy use of petroleum increase the budget, but it also emits more carbon [1]. Energy saving is one of the vital issues in today's world. To control the global economy, managing fuel consumption should be one of the main focus of ours. Although a lot of researchers have done significant work in this area.

The Texas A&M institute proposed that because of congestion, Americans have to travel more 5.5 billion hours and they need to buy an extra 2.9 billion gallons of fuel, which cost is about $121 billion.

An additional 56 billion pounds of Carbon Monoxide (CO) and greenhouse gas have been emitted for the congested conditions in 2011. The world is suffering from environmental pollution now [2, 3]. Reducing fuel consumption can also reduce carbon emission and sustain the clean and green atmosphere [4]. Although many researchers in the field of fuel and energy management have carried out significant research, the vehicle industry has also made several attempts to improve vehicle modernization for fuel efficiency and environmentally friendly technology that is economically viable [5, 6]. Vehicle speed prediction helps to control fuel management as well as reduce fuel consumption. If the speed of the vehicle can predict before the trip, then it would be a great contribution to energy optimization. In this research has predicted the speed of the vehicle based on different road status. A large number of datasets are collected from different sources. From one destination to another there will be multiple paths. Among these paths, the appropriate path has been predicted from the vehicle speed, which has consumed less energy and time-consuming.

In the track road curvature, speed-breaker, traffic, road condition, and weather condition will be considered to predict the estimated vehicle speed. Considering all of the situations, predicting the speed of the vehicle is quite difficult. In this research, the speed of the vehicle can be predicted for all situations. However, it is difficult to predict vehicle speed at the start of a trip, since road conditions can affect it [7]. But in this research, we can predict the speed of the vehicle before starting the ride by collecting the road situation.

Consumption of fuel is related to the road flow state. This state of flow is velocity related. If there is something on the road, prevents the speed of the vehicle and consumes more fuel. If the speed is at green velocity, then less fuel is consumed. More fuel consumption contributes to air pollution. If the track is chosen before starting the ride then more fuel consumption may be minimized. The fluidity of the road can be determined by estimating the speed whether it is congested or not. This also

**Corresponding author: kabir.ece07@gmail.com*

shows the time consumption by predicting speed and less time consumption is always preferable.

2. Literature Review

Many research has been done to predict the speed of the vehicle. Several techniques and technologies are employed to reduce fuel consumption to make the environment greener. Data-Driven Intelligent Transportation Systems (D2ITS) could be used to reduce fuel consumption that would clean and green the environment [8]. A Big Data-based Deep Learning Speed Prediction (BDDL-SP) algorithm is introduced capable of predicting vehicle speed in both the highway and urban networks [9]. There are different factors relating to vehicle speed such as driver behavior, route details, weather conditions, and traffic conditions are considered during the development of the data-driven model.

A 5-layer Adaptive Neuro-Fuzzy Inference System (ANFIS) is used to extract speed dynamics characteristics through supervised learning [10-12]. There are existing approaches can be classified into two categories. One is the Model-based method and another is the Data-based method [13]. The model-based approach is used mainly by vehicle dynamic driving force balance equation theory to create a prediction model between vehicle speed and different road conditions. Data-based methods approach linear regression models in [14-15], Kalman filter prediction models, [16-17], fuzzy rule prediction models [18-19], and neural network prediction models [20-22] has been proposed. Recently machine learning techniques, especially Neural Networks (NN) have been used to solve non-linear problems by using complex and multi-source field data. There has been a growing interest in NN's use in traffic research for decades, including driver behavior, parameter estimation, pattern analysis, traffic information forecasting, and so on [23]. An NN-based model was developed for estimating vehicle travel time, and the results showed that NN could be superior to conventional techniques [24]. However, some work has been done on traffic flow prediction. Stacked Auto-Encoders (SAE) was used to predict the flow of traffic using the Caltrans Performance Measurement System (CPMS) database [25]. Huang et al. used a deep belief net (DBN) in combination with Multitask Learning for traffic-flow prediction [26]. There are two steps to the work of predicting traffic flow-feature learning and model learning. In a recent study, a Long Short-Term Memory Neural Network (LSTM NN) was proposed for speed prediction with long time dependence, and the model can solve the problem of back-propagated error decay [27].

3. Methodology

Modern vehicles are designed basically to improve fuel consumption. Vehicle transition and congestion in the road are increasing day by day. The proposed route selection procedure for green fuel efficient uses different technologies. By observing the speed for a specific journey, the most fuel-efficient route can find out. This method shows us all the possible speed results for different conditions. If there is office time, traffic will be more. Then the speed of the vehicle will be less and that time fuel consumption will be more. Depending on the time, the traffic will be less or more. At noon, the traffic will be less. Even if the weather condition is good, then the speed of the car will be less. This will also affect the fuel consumption. But it will mostly depend on the driver's behavior.

In today's world, it is necessary to consume less amount of fuel so that it could be free from air pollution. Also, the global economy can be controlled. To solve fuel and time consumption, a machine learning model is established in this research.

By predicting the speed, how much fuel is needed to reach one destination to another can be known. A relationship between vehicle speed and fuel consumption has been given below.

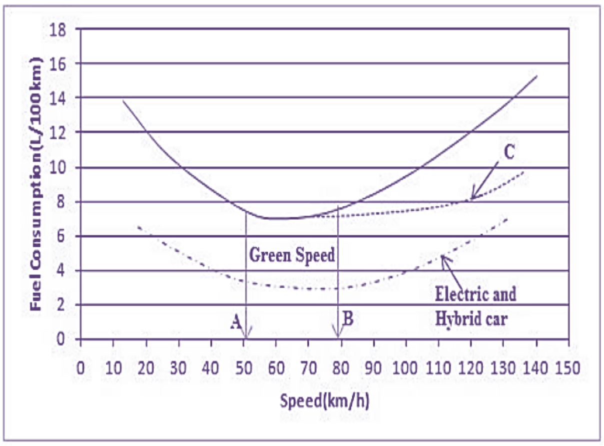

Fig. 1 - Relation between fuel consumption and speed [28].

From Fig. 1 it has shown that the vehicle runs above the green speed or below the green speed, then it will consume more fuel. The curve c is showing that if the aerodynamic drag is reduced at high speed, then it will consume less fuel [28].

Now to accomplish our research, the proposed algorithm is followed by data collection, data preparation, choose a model, train the model, evaluate the model, parameter tuning, make predictions.

3.1 Data Collection and Preparation

Machine learning requires a lot more data to train a model. The more data is collected, the result will more accurate. So the most important thing in machine learning is to gather data. Also, it depends on the quality and the quantity of data that describes how accurate our model is. The outcome of this step is the representation of the data which will use for training. Anyone can collect data manually and also some websites like Kaggle, UCI, etc. provide datasheets. The sample dataset of this research is given below.

Table 1 shows the data for bend or curvature conditions. In the first column, the vehicle is categorized in the car, truck, motorbike, etc. Next, for the curvature road how much the angle has to been shown in that column. Bend is divided into two categories. One is for the left bend and another is for the right bend. Distance value and road condition value is 0 here. Because this is used for another road status. Based on bend condition speed value is set here. If the angle is less like 10 degrees then the speed will be more which is 60 Km/h. And if the angle is much like 165 degrees, then the speed will be less which is 11 Km/h.

Table 1 - Dataset of vehicle speed for bend condition.

	vehicle	condition	distance	bend	road_condition	speed
0	car	left	0	10	0	60
1	car	left	0	15	0	58
2	car	left	0	20	0	56
3	car	left	0	25	0	54
4	car	left	0	30	0	52
...	...	...	...	...	...	...
275	motorbike	right	0	160	0	11
276	motorbike	right	0	165	0	11
277	motorbike	right	0	170	0	10
278	motorbike	right	0	175	0	10
279	motorbike	right	0	180	0	10

Table 2 - Dataset of vehicle speed for highway road condition.

	vehicle	condition	distance	bend	road_condition	speed
0	car	HighwayRoad	0	0	0	100
1	car	HighwayRoad	0	0	2	100
2	car	HighwayRoad	0	0	5	99
3	car	HighwayRoad	0	0	7	97
4	car	HighwayRoad	0	0	9	95
...	...	...	...	...	...	...
147	motorbike	HighwayRoad	0	0	67	23
148	motorbike	HighwayRoad	0	0	69	20
149	motorbike	HighwayRoad	0	0	71	17
150	motorbike	HighwayRoad	0	0	73	15
151	motorbike	HighwayRoad	0	0	75	10

Table 3 - Dataset of vehicle speed for urban road condition.

	vehicle	condition	distance	bend	road_condition	speed
0	car	UrbanRoad	0	0	0	50
1	car	UrbanRoad	0	0	2	50
2	car	UrbanRoad	0	0	5	50
3	car	UrbanRoad	0	0	7	49
4	car	UrbanRoad	0	0	9	49
...	...	...	...	...	...	...
147	motorbike	UrbanRoad	0	0	67	23
148	motorbike	UrbanRoad	0	0	69	20
149	motorbike	UrbanRoad	0	0	71	17
150	motorbike	UrbanRoad	0	0	73	15
151	motorbike	UrbanRoad	0	0	75	10

Table 2 shows the speed based on highway road conditions. On the highway, the highest speed is 100 Km/h. Road condition means that there can be holes or gravels on the road. Based on this, the road condition is categorized in different conditional values from 0 to 100. Where, 0 road conditions mean no holes, no speed breaker, and no congestion. On the other hand, read condition 100 means worse road condition. In Table 2 shows that, if the road condition is not good enough, then the value is defined as 50-75, and if the road condition is in good condition, then the value is defined as 0-20. Also if the road condition is medium, the value is defined 20-50 and the higher road condition, the lower vehicle speed and lower road condition mean higher vehicle speed.

Table 3 shows the speed of urban areas. In the urban areas, the maximum speed is about 50 Km/h. Just like the highway road, the condition of the road is categorized. If the condition of the road is not good enough, the value is specified as 50-75 and if the road conditions are in good condition, then the value is set as 0-20%. Also if the condition of the road is moderate, the value is 20-50%. if the read is worse, then the vehicle speed will be lower. In the case of better road conditions, then the vehicle speed will be higher. The vehicle speed will be moderate for moderate road condition.

Table 4 - Dataset of vehicle speed for the icy condition.

	vehicle	condition	distance	bend	road_condition	speed
0	car	icy	0	0	10	73
1	car	icy	0	0	13	71
2	car	icy	0	0	15	69
3	car	icy	0	0	17	67
4	car	icy	0	0	20	65
...	...	...	...	...	...	...
143	motorbike	icy	0	0	90	9
144	motorbike	icy	0	0	93	7
145	motorbike	icy	0	0	95	5
146	motorbike	icy	0	0	97	3
147	motorbike	icy	0	0	100	0

Table 5 - Dataset of vehicle speed for object condition.

	vehicle	condition	distance	bend	road_condition	speed
0	car	object	13	0	0	35
1	car	object	12	0	0	30
2	car	object	10	0	0	25
3	car	object	9	0	0	20
4	car	object	7	0	0	15
5	car	object	5	0	0	10
6	car	object	1	0	0	0
7	bus	object	13	0	0	35
8	bus	object	12	0	0	30
9	bus	object	10	0	0	25
10	bus	object	9	0	0	20
11	bus	object	7	0	0	15

Table 4 shows the speed based on the icy condition. In normal icy conditions, the maximum speed is about 73 Km/h. And in black icy conditions, the maximum speed is about 15 Km/h. Based on ice in the road; the value of road conditions is defined just like highway road conditions. If the road condition is not sufficiently good means too much ice on the road, then the value of the is defined as 70-100, etc. If the road conditions are in good enough, then the value is set within10-30. Even if the condition of the road is medium, the value is between 30-

70. Vehicle speed is lower for more slippery weather like 70-100 and higher for less slippery weather like 10-30.

If the object comes in front of the vehicle then what would be the speed. That is shown in Table 5 and Table 6. If the distance of the object is less, the speed will be small. On the other hand, if the distance of the object from the vehicle is large, the speed will be higher value.

Table 6 - Dataset of vehicle speed for speed breaker condition.

	vehicle	condition	distance	bend	road_condition	speed
0	car	SpeedBreaker	15	0	0	38
1	car	SpeedBreaker	13	0	0	35
2	car	SpeedBreaker	11	0	0	31
3	car	SpeedBreaker	9	0	0	28
4	car	SpeedBreaker	7	0	0	25
5	car	SpeedBreaker	5	0	0	20
6	car	SpeedBreaker	3	0	0	20
7	car	SpeedBreaker	1	0	0	18
8	bus	SpeedBreaker	13	0	0	35
9	bus	SpeedBreaker	11	0	0	31
10	bus	SpeedBreaker	10	0	0	30
11	bus	SpeedBreaker	8	0	0	26

Table 7 - Dataset of vehicle speed for local traffic condition.

	vehicle	condition	distance	bend	road_condition	speed
0	car	LocalTraffic(Red)	15	0	0	0
1	car	LocalTraffic(Red)	13	0	0	0
2	car	LocalTraffic(Red)	10	0	0	0
3	car	LocalTraffic(Red)	7	0	0	0
4	car	LocalTraffic(Red)	5	0	0	0
...	...	...	...	...	...	...
79	motorbike	LocalTraffic(Green)	10	0	0	30
80	motorbike	LocalTraffic(Green)	7	0	0	25
81	motorbike	LocalTraffic(Green)	5	0	0	20
82	motorbike	LocalTraffic(Green)	3	0	0	15
83	motorbike	LocalTraffic(Green)	0	0	0	0

On the local road, there have some speed-breaker. The number of speed-breaker can be one or more. Based on the distance of the speed breaker, the speed of the vehicle is defined. Just like the object condition speed is set.

There are also two types of traffic conditions. One is the Shared traffic area and another is the local traffic area. In the local traffic area, there will be vehicles like cars, trucks, etc. as like highways. For the local traffic area, the maximum speed is 10 Km/h. Just like the local traffic area, the dataset has been shown in Table 6 for the shared traffic area. In the shared traffic area, all types of vehicles will be available like a cycle, motorbike, car, etc. The maximum speed in the shared traffic area is 15 Km/h. There are three types of conditions in traffic, for example, green, yellow and red. In green, the vehicle can move. In yellow, it will start to move and in red, the vehicle won't be moved. Then the speed will be 0 Km/h which is represented in Table 7.

Table 8 - Dataset of vehicle speed for shared traffic condition.

	vehicle	condition	distance	bend	road_condition	speed
0	car	SharedTraffic(Red)	15	0	0	0
1	car	SharedTraffic(Red)	13	0	0	0
2	car	SharedTraffic(Red)	10	0	0	0
3	car	SharedTraffic(Red)	7	0	0	0
4	car	SharedTraffic(Red)	5	0	0	0
...	...	...	...	...	...	...
79	motorbike	SharedTraffic(Green)	10	0	0	6
80	motorbike	SharedTraffic(Green)	7	0	0	4
81	motorbike	SharedTraffic(Green)	5	0	0	2
82	motorbike	SharedTraffic(Green)	3	0	0	0
83	motorbike	SharedTraffic(Green)	0	0	0	0

All of the data is collected from different sources. Data are collected manually or there are some websites like Kaggle, UCI datasheet. Here the maximum speed of the vehicle is collected based on the speed limit of Australia for different road condition.

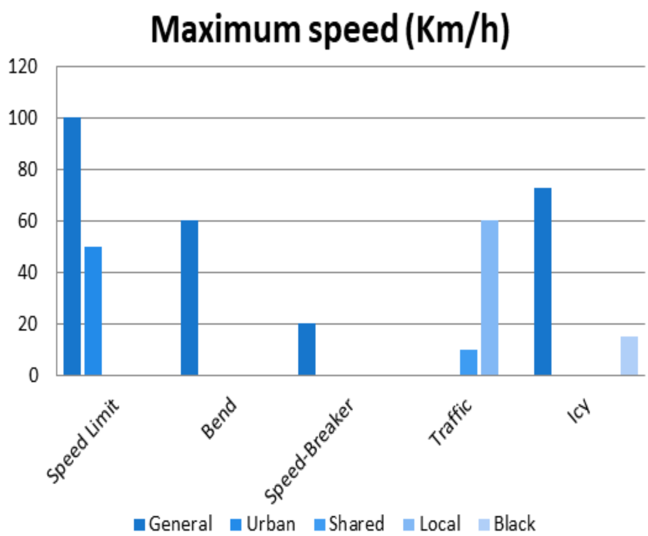

Fig. 2 - Maximum speed chart based on speed limits in Australia.

After collecting the data, we need to prepare data for training. Therefore while collecting data, some data can be missing or it can be duplicated. There may be also some errors in the datasheet. So these problems must need to solve before training. This raw data is not so helpful to train a model. The clean data have been required to remove duplicates, correct errors, dealing with missing values, normalization, data type conversions. The whole dataset was checked and the duplicate and missing values were removed and correct errors. Visualize data help to detect relevant relationships between variables or class imbalance. At last split dataset into training sets and test sets. The training set is used to train only and the test set is used for finding out accuracy and performance.

3.2 Choosing Algorithm

Now it's time to choose the right model. Many models can be used for different purposes. Selecting a model must-see does it full-fil the business goal. The model should be accurate, scalable, low time consuming and simple. If taking a complex model, it doesn't mean that it is a better model. Machine learning is one of the widely used data handling approaches. There are two parts to the machine learning algorithm. One is supervised learning and another is unsupervised learning. Apart from these, there is another algorithm which is called Reinforcement Learning. Commonly used some machine learning algorithms are Linear Regression, Logistic Regression, Decision Tree, K-means, Principal Component Analysis (PCA), Support Vector Machines (SVM), Naïve Bayes, Random Forest, and Neural Networks. To choose the right algorithm, find out the relation between input and output data. The plot of speed vs. bend, speed vs. distance and speed vs. road condition is given here. Classification of the machine learning algorithm is presented in Fig 4, 5, and 6.

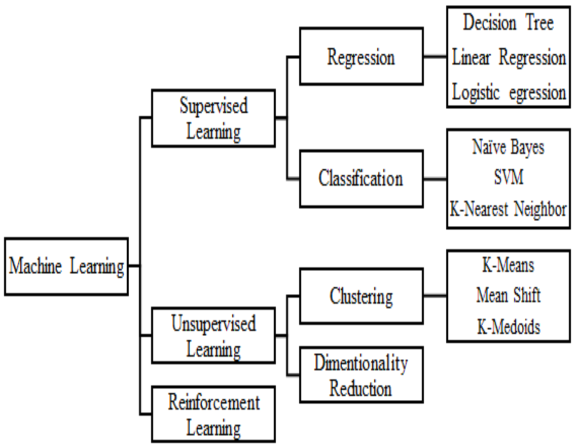

Fig. 3 - Machine learning algorithm classification.

In Fig. 4, we have seen that all the graphs are linear. In the speed vs. bend graph, one linear line is plotted based on one road condition. In the speed vs distance figure, four linear lines are plotted based on four different road conditions. And in speed vs. road condition, there is three plotted line. So, the linear regression algorithm has been chosen after observing the plot. This algorithm fits our model.

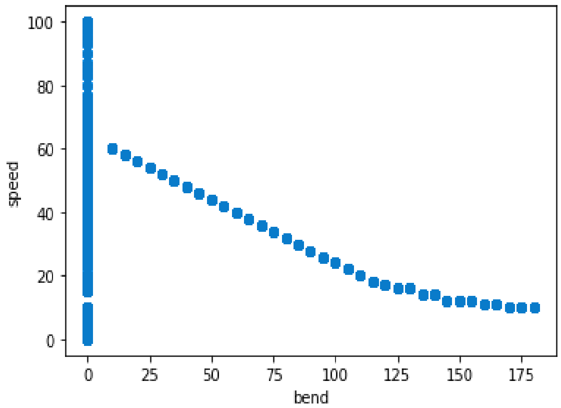

Fig. 4 - Graphical representation of speed vs bend.

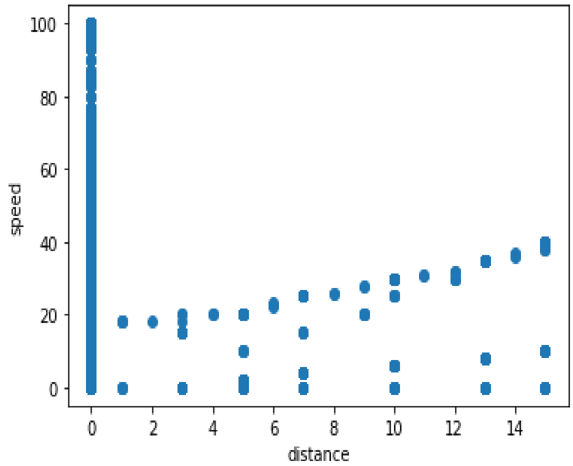

Fig. 5 - Graphical representation of speed vs object distance.

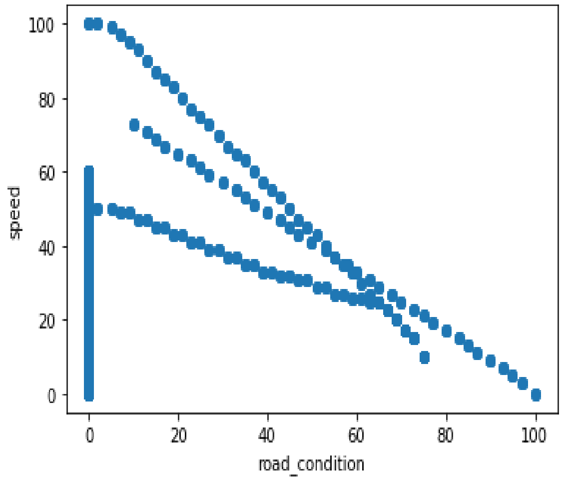

Fig. 6 - Graphical representation of speed vs road condition.

3.3 Equation of Linear Regression Algorithm

Avoid Linear regression is a linear model that assumes a linear relationship between the input variables and the single output variable. Let's

x = Input variable and

y = Output variable

More specifically, y can be measured from a linear combination of input variable x. When there are multiple input variables instead of a single input variable, then it can be referred to as multiple linear regression.

For any dataset,

$\{y_i, x_{i1}, \dots, x_{ip}\}^n i = 1$ of n statistical units

A linear regression model assumes the relationship between the input and the output variables. This relationship is modelled through a disturbance term or error variable ε.

Thus the model takes the form,

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \varepsilon_i = X_i^T + \varepsilon_i \quad (1)$$

$i = 1, \dots, n,$

Where T denotes the transpose, so the X_i^T is the inner product between X_i and β.

Often these n equations are stacked together and written in matrix notation as,

$$y = X\beta + \varepsilon \quad (2)$$

Where,

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

$$X = \begin{pmatrix} X_1^T \\ X_2^T \\ \vdots \\ X_n^T \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & & x_{1p} \\ 1 & x_{21} & \cdots & & x_{2p} \\ \vdots & \vdots & \ddots & & \vdots \\ 1 & x_{n1} & & \cdots & x_{np} \end{pmatrix},$$

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}, \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Graphical representation of Linear Regression is given below

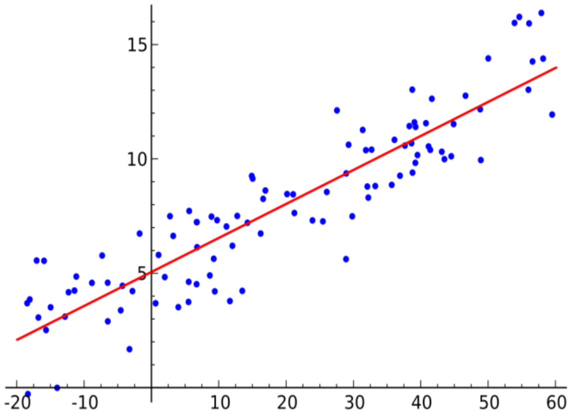

Fig. 7 - The plot of the linear regression model.

Fig-7 shows the relation between x and y-axis. With the increasing of x, y will also increase. The straight line indicates the maximum points on the best match line.

3.4 Training and Evaluating the Model

Now training a model is the most important thing of machine learning. After cleaning out the dataset, the linear regression model is used to train up the model. As a programming language Python is used here. First, the dataset was exported to the Jupiter notebook. Then has created a dummy variable of the vehicle and condition column, because machine learning algorithm only works with numerical values.

Table 9 - Creation of proposed dummy variables.

bus	car	motorbike	truck	...	OtherRoad	SharedTraffic(Green)	SharedTraffic(Red)	SharedTraffic(Yellow)	SpeedBreaker	UrbanRoad	icy	left	object	right
0	1	0	0	...	0	0	0	0	0	0	0	1	0	0
0	1	0	0	...	0	0	0	0	0	0	0	1	0	0
0	1	0	0	...	0	0	0	0	0	0	0	1	0	0
0	1	0	0	...	0	0	0	0	0	0	0	1	0	0
0	1	0	0	...	0	0	0	0	0	0	0	1	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
0	0	1	0	...	0	0	0	0	0	1	0	0	0	0
0	0	1	0	...	0	0	0	0	0	1	0	0	0	0
0	0	1	0	...	0	0	0	0	0	1	0	0	0	0
0	0	1	0	...	0	0	0	0	0	1	0	0	0	0

Then split the dataset into the input and the output column. After that linear regression function is called to train up the dataset. To predict the result and accuracy, the dataset is split into two-sector. One section is called a training dataset and another is called the testing dataset. In supervised machine learning, the model is built with labeled sample data and in unsupervised machine learning tries to draw inference from non-labeled data.

The model is being prepared and now the time to check it out. It will be used to test its performance from unused data, which is called test data. This will represent how the model works in the real world. A good rule for splitting out the dataset is 80/20 or 70/30. Most of this depends on the size of the source. For a lot of data, don't need to take a big fraction for evaluating the dataset. Here is given the entire process in graphical representation.

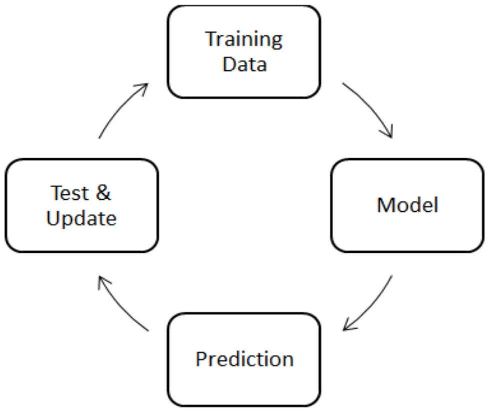

Fig. 8 - Graphical representation of evaluating the model.

3.5 Tuning Parameter

After evaluating the model, this section refers to the tuning of parameters to improve the model. If there are any chances of improving the model, just need to see. There were a few parameters implicitly assumed while training and now test those assumptions and try other values. Increasing the number of training cycles can lead to more accurate results.

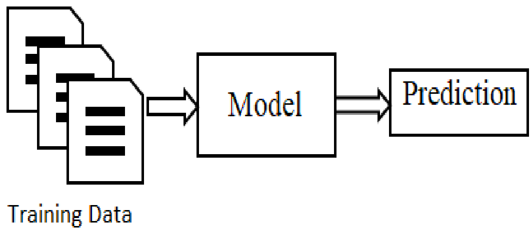

Fig. 9. Graphical representation of tuning parameter.

The model has shown the performance and accuracy results. The predicted result has found 23.72844208 for the label data 23. For all testing data, the accuracy of the proposed algorithm is found around 95%.

The steps of machine learning are showing graphically below.

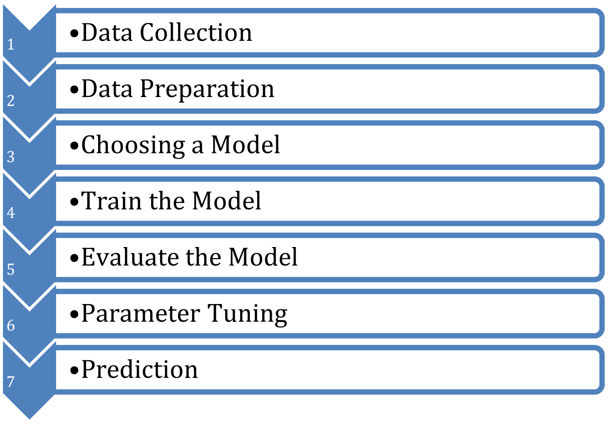

Fig. 10- Graphical representation of machine learning steps.

4. Result and Discussion

The fitted model gives accuracy and performance results based on our dataset. The model has given the predicted speed of the vehicle and successfully it has shown the approximate result of the vehicle which is the nearest speed value of the test dataset. To predict a result a number of array is given to test which is as [0,0,51,0,1,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0]. This is the sample input data matrix of the proposed algorithm.

Table 10 - The output of predicted and actual result.

Input	Predicted Result	Actual Result
[0,0,51,0,1,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0]	23.72844208	23
[7,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0]	19.61171628	15
[0,40,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1]	48.36267629	48
[0,0,40,0,0,1,0,0,0,0,0,0,0,0,0,1,0,0,0]	49.4656284	49
[0,0,73,0,0,1,0,0,0,0,0,0,0,0,0,1,0,0,0]	22.08126224	23
[0,0,39,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,0]	33.67277412	33
[0,90,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,1]	32.00253213	28
[0,140,0,0,0,0,1,0,0,0,0,0,0,0,0,0,1,0,0]	13.78513371	14
[0,70,0,0,0,1,0,0,0,0,0,0,0,0,0,0,1,0,0]	37.13369396	36
[0,0,49,0,0,0,1,0,0,0,1,0,0,0,0,0,0,0,0,0]	49.26921818	45

The accuracy of the proposed algorithm is 95%. Therefore the performance is fully dependent on large amounts of datasets. The more data is given to the model, the more it will learn. The performance will be much better than before.

In the above, fuel consumption based on different road conditions has been shown. In the free-flow condition, route 2 is fuel-efficient. In a moderate congestion condition, the road is tolerable and traffic density is at random nature. This time is more fuel-efficient. But in congested road conditions, the road is very busy, especially in-office hours. At that time more fuel is consumed [28].

From above, fuel consumption is related to the flow condition of the road. This flow condition is related to speed. If the speed is less, then there is something in the road that is preventing the vehicle speed and consume more fuel. If the speed is in the green speed, then less amount of fuel has been consumed. More fuel consumption makes more air pollution. So before starting the trip, if the track is selected appropriately then it can reduce more fuel consumption. By predicting the speed, the fluency of the road can be known whether it is congested or not.

Table 11 - Free flow condition fuel consumption [28].

Performance Measure	Route 1	Route 2	Route 3	Remarks
Distance (Km)	12.1	10.8	11.2	
Running time (Minutes)	12 m	11 m	12 m	
Stop time (Minutes)	2 m	2 m	2 m	
Total time (Minutes)	14 m	13 m	14 m	Assumption-1
Total distance w.r.t time	14 km	13 Km	14 Km	
Fuel used (Liter)	1.82	1.69	1.456	
Fuel Consumption (Lt/Km)	0.13	0.13	0.13	

Table 12 - Performance on moderate congested road condition [28].

Performance Measure	Route 1	Route 2	Route 3	Remarks
Distance (Km)	12.1	10.8	11.2	
Running time (Minutes)	17 m	18 m	18 m	
Stop time (Minutes)	4 m	4.5 m	4 m	
Total time (Minutes)	21 m	22.5 m	22 m	Assumption-2
Total distance w.r.t time	21 km	22.5 Km	22 Km	
Fuel used (Liter)	2.73	2.925	2.86	
Fuel Consumption (Lt/Km)	0.13	0.13	0.13	

Table 13 - Performance on heavily congested road condition [28].

Performance Measure	Route 1	Route 2	Route 3	Remarks
Distance (Km)	12.1	10.8	11.2	
Running time (Minutes)	20 m	21 m	18 m	
Stop time (Minutes)	8 m	9 m	8 m	
Total time (Minutes)	28 m	30 m	26 m	Assumption-2
Total distance w.r.t time	28 km	30 Km	26 Km	
Fuel used (Liter)	3.64	3.9	3.38	
Fuel Consumption (Lt/Km)	0.13	0.13	0.13	

From above, the fuel consumption is related to the flow condition of the road. This flow condition is related to speed.

In paper [20], the development of D2ITS was discussed and several important components of D2ITS were introduced, including vision, multi-source, and learning-driven Intelligent Transportation System (ITS). They used two prediction models based on neural networks. One is Time-Based Neural Networks and another is Distance-Based Neural Network. They only predict the speed of the vehicle for specific urban driving.

A short-term velocity predictor for vehicles is developed in this paper [29]. They used the fuzzy Markov model and the Auto-regressive model. This experiment was tested in urban areas for short term traffic speed prediction.

This paper proposed for short-term traffic speed information prediction [30]. In this paper, the Deep Belief

Network (DBN) model is used. DBN's output indicates it is successful in the prediction field of traffic information.

The aim of this paper [7] is to investigate the application of Deep Learning techniques to this problem to determine the driver-specific vehicle speed profile for an individual driver's repeated drive cycle at the beginning of a drive cycle, which can be used in an optimization algorithm to minimize the amount of fossil fuel energy used during the journey [7]. The deep learning algorithm is used here to predict the speed. This Neural Network, which uses real-time route data at the beginning of the ride, improves accuracy over the direct use of the real-time data but does not improve accuracy if the speed limit is used directly.

In paper [9] proposed speed prediction algorithm using the actual driving data collected by one test drive. Experiment result indicates that the algorithm is capable of predicting the speed in highway and urban traffic networks [9]. BDDL-SP algorithm is used here. Only predicted the result in driver behavior, route type, traffic and weather condition.

In comparison, our paper proposed a vehicle speed prediction model predicts vehicle speed for all the conditions. Machine learning is used here and the model is trained using a linear regression algorithm which is simple in computation and efficient. However, for more accurate performance, more real-time data is needed to train up the model.

5. Conclusion

In this paper, we have proposed a machine learning model, which is used to find out the speed of the vehicle in various factors such as route type, traffic condition, road condition, weather condition and also in road curvature. Large data has been collected from different sources. This data is used to train up the model by applying it into the linear regression model. All the experiment was done by using the python programming language. The predicted value and the actual value were found almost the same after training up the model. From this model, the accuracy has found 95%. In the future, we will try to find out more accurate results with more accurate speed prediction by collecting more relevant data using a reinforcement learning algorithm.

References

[1] Mahlia, T. M. I., Tohno, S., & Tezuka, T. (2013). International experience on incentive program in support of fuel economy standards and labelling for motor vehicle: A comprehensive review. Renewable and Sustainable Energy Reviews, 25, 18-33.

[2] Mahlia, T. M. I., Tohno, S., & Tezuka, T. (2012). History and current status of the motor vehicle energy labeling and its implementation possibilities in Malaysia. Renewable and Sustainable Energy Reviews, 16(4), 1828-1844.

[3] Uddin, M. N., Daud, W. W., & Abbas, H. F. (2013). Potential hydrogen and non-condensable gases production from biomass pyrolysis: insights into the process variables. Renewable and Sustainable Energy Reviews, 27, 204-224.

[4] Salvi, B. L., Subramanian, K. A., & Panwar, N. L. (2013). Alternative fuels for transportation vehicles: a technical review. Renewable and Sustainable Energy Reviews, 25, 404-419.

[5] Deakin, E. (2001). Sustainable development and sustainable transportation: strategies for economic prosperity, environmental quality, and equity.

[6] Uddin, M. N., Daud, W. W., & Abbas, H. F. (2014). Effects of pyrolysis parameters on hydrogen formations from biomass: a review. Rsc Advances, 4(21), 10467-10490.

[7] Lemieux, J., & Ma, Y. (2015, October). Vehicle speed prediction using deep learning. In 2015 IEEE Vehicle Power and Propulsion Conference (VPPC) (pp. 1-5). IEEE.

[8] Alsabaan, M., Alasmary, W., Albasir, A., & Naik, K. (2012). Vehicular networks for a greener environment: A survey. IEEE Communications Surveys & Tutorials, 15(3), 1372-1388.

[9] Cheng, Z., Chow, M. Y., Jung, D., & Jeon, J. (2017, June). A big data based deep learning approach for vehicle speed prediction. In 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE) (pp. 389-394). IEEE.

[10] Goode, P. V., & Chow, M. Y. (1994, June). A hybrid fuzzy/neural system used to extract heuristic knowledge from a fault detection problem. In Proceedings of 1994 IEEE 3rd International Fuzzy Systems Conference (pp. 1731-1736). IEEE.

[11] Chow, M. Y., Altug, S., & Trussell, H. J. (1999). Heuristic constraints enforcement for training of and knowledge extraction from a fuzzy/neural architecture. I. Foundation. IEEE Transactions on Fuzzy Systems, 7(2), 143-150.

[12] Altug, S., Chow, M. Y., & Trussell, H. J. (1999). Heuristic constraints enforcement for training of and rule extraction from a fuzzy/neural architecture. II. Implementation and application. IEEE Transactions on Fuzzy Systems, 7(2), 151-159.

[13] Mori, U., Mendiburu, A., Álvarez, M., & Lozano, J. A. (2015). A review of travel time estimation and forecasting for advanced traveller information systems. Transportmetrica A: Transport Science, 11(2), 119-157.

[14] Zhang, P., Yan, F., & Du, C. (2015). A comprehensive analysis of energy management strategies for hybrid electric vehicles based on bibliometrics. Renewable and Sustainable Energy Reviews, 48, 88-104.

[15] Liu, S., Wang, J., Zhang, L., & Fang, S. (2015). Characteristic and prediction model for operating speed model in urban underground road.

[16] Chunbo, X., Xiao, R., & Yanqing, L. (2014). Short-term prediction method of wind speed series based on Kalman filtering fusion. Transactions of china electrotechnical society, 29(2), 253-259.

[17] Yang, G., XV, R., Qing, M., Zheng, K. L., & Zhang, B. (2017). Short-term traffic volume forecasting based on ARMA and kalman filter. J. Zhengzhou Univ.(Eng. Sci.), 38, 36-40.

[18] Xie, H. (2014). Prediction of driving condition for Plug-in Hybrid Electric Vehicles. Chongqing: College of mechanical engineering of Chongqing University.

[19] Xie, S., Jiang, X., & Wei, L. (2014). Prediction of Driving Speed Interval Passenger Car Based on Fuzzy Linear Regression. [J]. Journal of Highway and Transportation Research and Development, (08), 126-130.

[20] Zhang, L., Zhang, J., & Liu, D. Neural Network based Vehicle Speed Prediction for Specific Urban Driving. In 2018 Chinese Automation Congress (CAC) (pp. 1798-1803). IEEE.

[21] Yuan, L. (2016). Vehicle Speed Prediction and Application Based on NAR Neural Network. [D]. DALIAN UNIVERSITY of TECHNOLOGY.

[22] Zhao, X. H., Yao, Y., Wu, Y. P., Chen, C., & Rong, J. (2016). Prediction model of driving energy consumption based on PCA and BP network. [J]. Journal of transportation Systems Engineering and Information Technology, (05), 185-191.

[23] Dougherty, M. (1995). A review of neural networks applied to transport. Transportation Research Part C: Emerging Technologies, 3(4), 247-260.

[24] Hua, J., & Faghri, A. (1994). Apphcations of artificial neural networks to intelligent vehicle-highway systems. Transportation Research Record, 1453, 83.

[25] Lv, Y., Duan, Y., Kang, W., Li, Z., & Wang, F. Y. (2014). Traffic flow prediction with big data: a deep learning approach. IEEE Transactions on Intelligent Transportation Systems, 16(2), 865-873.

[26] Huang, W., Song, G., Hong, H., & Xie, K. (2014). Deep architecture for traffic flow prediction: deep belief networks with multitask learning. IEEE Transactions on Intelligent Transportation Systems, 15(5), 2191-2201.

[27] Ma, X., Tao, Z., Wang, Y., Yu, H., & Wang, Y. (2015). Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. Transportation Research Part C: Emerging Technologies, 54, 187-197.

[28] Nasir, M. K., Md Noor, R., Kalam, M. A., & Masum, B. M. (2014). Reduction of fuel consumption and exhaust pollutant using intelligent transport systems. The Scientific World Journal, 2014.

[29] Jing, J., Filev, D., Kurt, A., Özatay, E., Michelini, J., & Özgüner, Ü. (2017, June). Vehicle speed prediction using a cooperative method of fuzzy Markov model and auto-regressive model. In 2017 IEEE Intelligent Vehicles Symposium (IV) (pp. 881-886). IEEE.

[30] Jia, Y., Wu, J., & Du, Y. (2016, November). Traffic speed prediction using deep learning method. In 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC) (pp. 1217-1222). IEEE.