

Analysis on Used Car Dataset using Machine Learning Models

Prerna Paliwal, Charul Singhvi, Aditya Maheshwari

Abstract: The research paper focuses on study of used cars of different models based on different fuel types, owner types and years all at different locations and also other factors like Mileage, Engine type, Power consumed and number of seats available. Data is visualized on the basis of Kilometers driven, Fuel Type and Owner Type.

Keywords: Machine Learning, Models, Fule Types, Owner Type, Engine Capacity, Kilometers driven

I. INTRODUCTION

Used cars are mainly the second hand, third handed or furthermore handed cars. The buyers look for the efficiency of the car in the form of the kilometers driven by the car and the mileage it is giving. Many factors affect the price of the used cars but these are major factors mentioned above majorly affect the efficiency of used cars. The price is calculated mainly by considering the factors such as fuel type, transmission type, model, company, location also affect the price of the cars.

In this study the above-mentioned constraints are visualized and accordingly the neural networks [1] are created and the new price is predicted on the basis of the old price and the other factors such as kilometers driven, engine capacity [2], power consumed [3], transmission type, year, location, mileage [4] and owner type.

II. METHODOLOGY

The dataset is taken from Kaggle <https://www.kaggle.com/avikasliwal/used-cars-price-prediction> The dataset includes 6019 instances each row having 14 attributes for differentiation. There are 1876 unique values of different car models. It includes data of 11 different locations and also, of years ranging from 1988 to 2019 having fuel types petrol, diesel, LPG, CNG, electric.

A. Figures and Tables

The figures depict the sub grouped data into 2 categories: Automatic and Manual and both the categories consist of various fuel types. In this plot we are comparing the used car price with the new car price on the basis of kilometers driven by the car, mileage and fuel type.

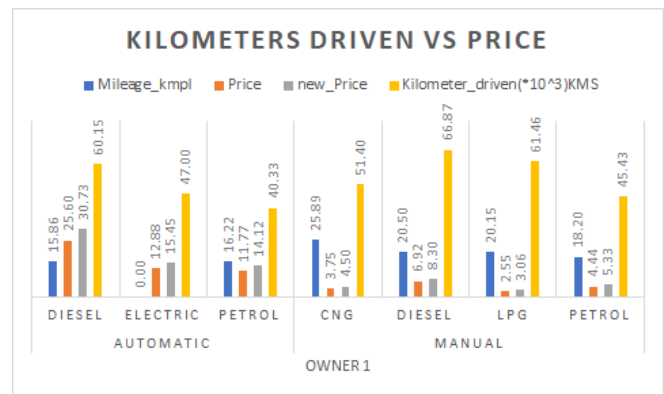


Fig. 1. Kilometers driven v/s Price of a car for First Owner.

From this plot (Fig. 1) we conclude that less is the number of kilometers driven by the car higher is the efficiency of the car and higher is the cost price for the First Owner .

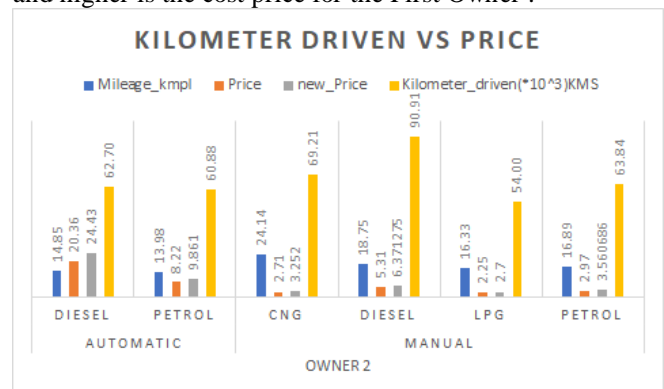


Fig. 2. Kilometers driven v/s Price of a car for Owner Second

From this plot (Fig. 2) we conclude that less is the number of kilometers driven by the car higher is the efficiency of the car and higher is the cost price for the Second Owner.

Revised Manuscript Received on March 5, 2020.

Prerna Paliwal*, Techno India NJR Institute of Technology, Udaipur 313001

prernapaliwal80@gmail.com*

Charul Singhvi, Techno India NJR Institute of Technology, Udaipur 313001

Aditya Maheshwari, Techno India NJR Institute of Technology, Udaipur 313001

Table- 1: Used Cars data for first and second owners

Owner Type	Transmission	Fuel Type	Mileage (in kmpl)	Price (in lakh)	New Price (in lakh)	Kilometers Driven (in 10 ³ kms)
First	Automatic	Diesel	15.86	25.6	30.73	60.15
		Electric	0	12.88	15.45	47
		Petrol	16.22	11.77	14.12	40.33
	Manual	CNG	25.89	3.75	4.5	51.4
		Diesel	20.5	6.92	8.3	66.87
		LPG	20.15	2.55	3.06	61.46
Second	Automatic	Petrol	18.2	4.44	5.33	45.43
		Diesel	14.85	20.36	24.43	62.7
		Petrol	13.98	8.22	9.86	60.88
	Manual	CNG	24.14	2.71	3.25	69.21
		Diesel	18.75	5.31	6.37	90.91
		LPG	16.33	2.25	2.7	54
		Petrol	16.89	2.97	3.56	63.84

Table- II includes the data grouped on the basis of Owners and sub grouped on the basis of Transmission types. This table justifies the relation between different features of a car to compare the price of used cars and new cars.

III. RESULT AND DISCUSSION

Predicted by Observed

Target: Price

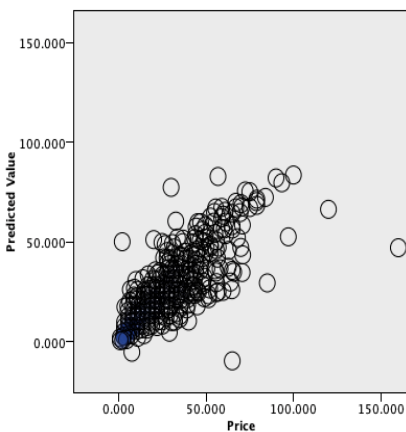


Fig. 3. shows the relation between Predicted value and original value of Old Price with respect to dataset.

Predicted by Observed

Target: new_Price

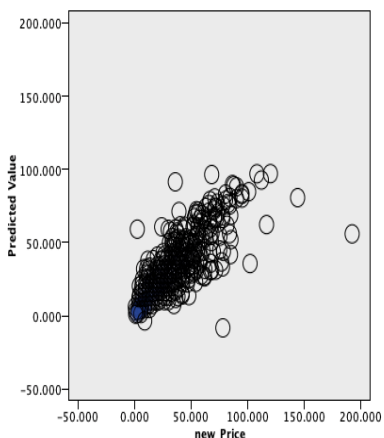


Fig. 4. shows the relation between Predicted value and original value of New Price with respect to dataset.

Network

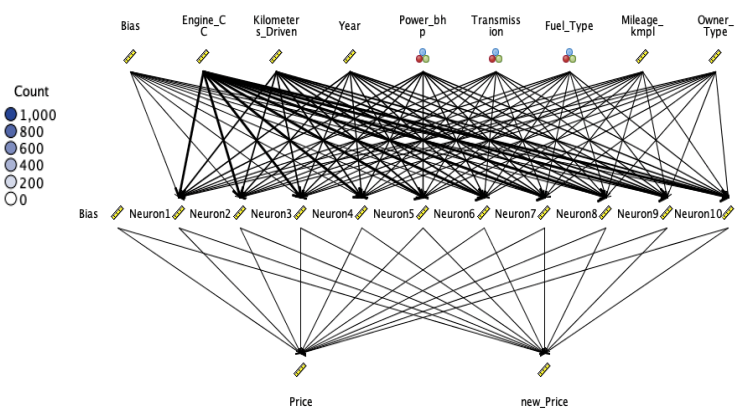


Fig. 5. shows the Neural Network to Predict the New Price. This network consists of 10 Hidden Layers. These networks can learn and model the relationships between inputs and outputs that are complex and nonlinear.

Field	Sample Graph	Measurement	Min	Max	Mean	Correlation
Kilometers_D...		Continuous	171.000	6500000...	58738.3...	-0.011
Year		Continuous	1998.000	2019.000	2013.358	0.305
Mileage_kmpl		Continuous	0.000	33.540	18.135	-0.307
Price		Continuous	0.440	160.000	9.479	--

Fig. 6. shows the relation of each attribute (Kilometers Driven, Year, Mileage (in kmpl), Price (in lakh)) with target value (New Price) and also plot of each attribute.

Model Summary

Targets	Price new_Price
Model	Multilayer Perceptron
Stopping Rule Used	Error cannot be further decreased
Hidden Layer 1 Neurons	10

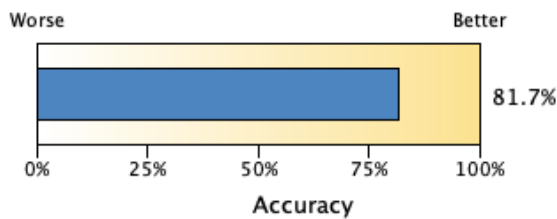


Fig. 7. consists of Summary of Model which consist of Accuracy of Model which is 81.7% and define the target of model uses Multilayer Perceptron Class of Neural Network to find the relation between attributes and target value.

IV. CONCLUSION

Through this study we conclude that as the efficiency of a car directly depends on the owner type that is whether it is the first, second, third or above owner. As the owner type deprecates the price depreciates and also as the kilometers driven increases the price of the car decreases.

REFERENCES

1. Rose, D. (2003) "Predicting Car Production using a Neural Network Technical Paper- Vetronics (In-house)". Thesis, U.S. Army Tank Automotive Research, Development and Engineering Center (TARDEC).R. Rojas: Neural Networks, Springer-Verlag, Berlin, 1996.
2. <https://carbiketech.com/engine-capacity-cc/>
3. http://en.citizendium.org/wiki/Energy_consumption_of_cars
4. Mileage: <https://en.wikipedia.org/wiki/Mileage>.

AUTHOR PROFILE



Prerna Paliwal is pursuing B. Tech in Computer Science and Engineering from Techno India NJR Institute of Technology Udaipur. Her research interest is Machine Learning and Statistical Modelling.



Charul Singhvi is pursuing B. Tech in Computer Science and Engineering from Techno India NJR Institute of Technology Udaipur. Her research interest is Neural Network and Tensor Flow.



Aditya Maheshwari He is working in the capacity of Project Lead- New Initiative & Research at Techno India NJR Institute of Technology Udaipur.