Pakistan Journal of Engineering and Technology, PakJET
Multidisciplinary & Peer Reviewed
Volume: 4, Number: 2, Pages: 113- 119, Year: 2021

Used Cars Price Prediction using Machine Learning with Optimal Features

Muhammad Asghar[1], Khalid Mehmood[2], Samina Yasin[3], Zimal Mehboob Khan[4]

[1]University of South Asia Lahore, Pakistan
[2]Govt. Associate College for Women, Rahim Yar Khan, Pakistan
[3]University of Agricultural Faisalabad, Pakistan
[4]Khwaja Freed University of Engineering and Information Technology, Rahimyar Khan, Pakistan

Corresponding author: Muhammad Asghar (email: raoasghar777@gmail.com)

Abstract:- **We all are need the personal vehicle that could help us to travel from home to office and travel to vocations means we need the personal vehicle for traveling for this we purchase the new vehicle or used vehicle this is some time take so much to take decision for purchasing the new one and most difficult decision is to take how to sale the old one that is already we have keep using if we sale and what is best price we can get or gives us more benefits. More over the purchasing power of the customers is low due to the prices of the new cars. There are different methods to predict the price of the car according to market value. Our proposed method helps the both the purchase and seller for to purchase and sale their vehicle and they can predict the best for their vehicle and make their decision good for personal and business. Our proposed model performance shows that the proposed study is productive and efficient. In the proposed study the machine learning algorithm Regression helps in the outperform. Here we use the Statistical test to get the design value of P and get the optimal features and using the linear regression. First, we find the RFE and then apply the statistical test for VIF for the OLS Regression. Prediction results shows the study is efficient and effective.**

Index Term-- **Car price Prediction, Machine learning, Regression, Statistical test, IF, OLS.**

I. INTRODUCTION

The Vehicle are in the different categories like Car, Sedan, coupe, supports car, station wagon, hatchback, convertible, sport-utility vehicle (SUV) minivan, Pickup truck, each vehicle having the different features and type of uses based on the features the prices of the vehicles are change. Here are some features we are discussing that are common in the vehicles like, Odometer, 4-wheel drive, and 2-wheel drive, Transmission type and mileage, these are the features peoples are consider as they see in the market move but the actual facts what we find in the proposed study [1]. Here we have considered three type of the target peoples one is who are the seller they only sell the cars we call them dealers. They are one of the most important target groups that may be interested in the study's findings. Used car dealers that have a better understanding of what makes a car appealing and what the most significant qualities are for a used car will be able to apply this knowledge and provide better service. many people have been interested in the used car market at some time in their lives because they wanted to sell or acquire a used car. It's a great mistake to pay too much or sell for less than the market value in this procedure. there are websites that proved an estimate of a car's worth [2]. They could have an excellent model for forecasting. Having a second model, on the other hand, may aid them in providing a better forecast to their consumers. As a result, the model established in this study might aid online web services that determine the market worth of a used car [3].

II. LITEARE REVIEW

Car pricing using machine learning has a clear correlation with the information acquisition process for technical systems. The purchase or sale on Internet market websites has recently become the primary technique for knowledge acquisition [4]. We may divide it into two forms after finding the data: organized and unstructured, involving knowledge context extraction, data inference, and qualitative data laws [5]. The key purpose of the current analysis is to investigate various forms of car data and to establish an integrated met forecasting car data. time-consuming method of recommendation, posting for car-based analysis. This paper would discuss the methods of methodology for forecasting car data [1]. We create a deterministic inventory model for an item in this work whose demand depends on both the sale price and the period after the last replenishment of the we believe that the demand rate incorporates the results of sale expense and a time function in an additive way. In addition, we assume that the cost of keeping is a power function of the amount of time a form keeps product period and the sale price that increase the overall unit time of inventory benefit. To solve this inventory problem, we present an effective algorithm [6]. Any numerical examples are given to demonstrate how the algorithm works [7]. More than ever, cars are being sold. Instead of renting a new vehicle because of cost, developed nations follow the leasing community. The increase in sales of used cars is, thus, rising exponentially. By listing unreasonable prices because of the demand, car dealers often take advantage of this condition [8].

Therefore, a need exists for a model that can attach a price to a vehicle by analyzing its attributes, considering the pr we use the Random Forest approach of supervised learning to forecast the prices of used vehicles [9]. After diligent exploratory data analysis, the model was selected to assess the effect of each feature on the price. A Random Forest was created to train the data with 500 Decision Trees. The teaching accuracy was found to be 95.82 percent from experimental findings and the accuracy of the testing was 83.63 percent. The model will reliably forecast the price of cars by selecting the most associated characteristics of other vehicles [2]. This research presents a new technique to address the car price occupancy rate prediction based on Deep Learning with Recurrent Neural Networks [1]. In smart mobility, this is an interesting problem and we approach it in an innovative way here, consisting of automatically designing a deep network that encapsulates the behavior of the occupancy of the car and then being able to make an informed guess about the number of free parking spaces near the medium-term horizon [10]. We analyze a real-world case study consisting of the occupancy values of 29 cars in Birmingham, UK, during eleven weeks and compare our results to other predictors in the state-of-the-art. The findings show that our strategy is precise to the extent that it is useful to be used by citizens in their daily lives, as well as to outperform existing competitors [3].

The rapid growth of internet shopping and e-commerce websites, e.g., eBay and OLX, has been seen in recent years. Any day, online shopping markets deliver millions of items for sale. Such commodities are classified into several types of items. For vendors, it is important to determine the price of the second-hand item accurately. The price of only one object type can be predicted by state-of-the-art techniques. In addition, the price range of a given second-hand object in the prediction challenge was not used by any of the current approaches, because there are many ads for the same good at different prices. In this vein, we suggest deep model architecture for forecasting the price as the first contribution [11]. This suggested methodology uses a deep neural network for price prediction that involves long short-term memory (LSTM) and convolution neural network architectures.

In comparison with the support vector machine baseline model, the proposed model obtained a higher mean absolute error accuracy performance [12]. In comparison, two inputs are used in the second contribution [10]. Next, we recommend the forecasting of the second-hand item's minimum and maximum rates. Linear regression, LSTM and seasonal autoregressive integrated moving average approaches are included in the models used for the forecasting mission. Second, in forecasting the item quality score, we suggest using the model of the first contribution. Then, the commodity quality score and the minimum and maximum prices forecast are combined to provide the final expected price of the item. Using a dataset for second-hand goods from a database, the proposed approach of comparing the projected second-hand item quality score with the predicted minimum and maximum price outperforms the other models with a large output difference in all the accuracy metrics used [13, 14].

Using three cross-sections of pricing data from U.S. airport markets covering the years 2005 to 2016, we analyze the price impact of restructuring in the car rental industry. During this time, the car rental industry went through a series of mergers, contributing to a significant rise in market concentration. We find that ownership concentration has a different impact on the market (weekday) and leisure (weekend) divisions. With the rise in business saturation, average weekday prices increased by 2.1 percent and weekend prices dropped by 3.3 percent. In view of the seasonal fluctuations in demand from business and leisure passengers, given the occasional demand disparities between business and leisure travelers, we describe this result with a horizontal product differentiation model that facilitates variation in the categories of consumers and the marginal costs of businesses [15-18].

Consolidation results in marginal cost savings, but the degree to which these savings are transmitted to multiple types of consumers depends on the extent of the cost of swapping. In fact, higher rates are paid to weekday consumers with high transition costs due to the expanded pricing control of vendors, while the more price-sensitive weekend segment enjoys lower prices facilitated by productivity gains. Our findings illustrate that restructuring may have distinct welfare impacts on multiple consumer segments, and merger studies can account for the heterogeneous influence focused on the practices of price discrimination of firms rather than only taking into account average effects[18-25].The next section is the data collection and this is the actually part of proposed study methodology.

III. DATA COLLECTION

The Data is collected from Kaggle and used for the purposed study we have attributes in the data are given below. Here some are important and some are not important so we need cleaning processes and in the cleaning processes we remove the some of the attributes are only keep the most important and strong correlation attributes. *Car_id, peakrp , carName* this kind of attributes are need to remove and the remaining are also filter through the processes of the analysis and then cleaning of the attributes.

Figure 1 shows the view of the dataset and data type and name and number of attributes here we have 25 attributes but the actual we use after cleaning processes are will be discussed in the next heading.

#	Column	Non-Null Count	Dtype
0	car_ID	205 non-null	int64
1	symboling	205 non-null	int64
2	CarName	205 non-null	object
3	fueltype	205 non-null	object
4	aspiration	205 non-null	object
5	doornumber	205 non-null	object
6	carbody	205 non-null	object
7	drivewheel	205 non-null	object
8	enginelocation	205 non-null	object
9	wheelbase	205 non-null	float64
10	carlength	205 non-null	float64
11	carwidth	205 non-null	float64
12	carheight	205 non-null	float64
13	curbweight	205 non-null	int64
14	enginetype	205 non-null	object
15	cylindernumber	205 non-null	object
16	enginesize	205 non-null	int64
17	fuelsystem	205 non-null	object
18	boreratio	205 non-null	float64
19	stroke	205 non-null	float64
20	compressionratio	205 non-null	float64
21	horsepower	205 non-null	int64
22	peakrpm	205 non-null	int64
23	citympg	205 non-null	int64
24	highwaympg	205 non-null	int64
25	price	205 non-null	float64

FIGURE 1: The dataset attributes and its datatypes.

IV. DATA EXPLORATION

After cleaning the data, the ready for exploring in Fig. 2 the prices related to dataset the figure shows that the mostly vehicles are falling in range of $10K to $20K and this shows that that mostly market values vehicles are in this range of prices the figure also shows us the most vehicles are having the price of $10K in the overall dataset.

It means the expansive vehicle are not mostly used vehicles they are new and people use them and not to go for purchase a used or old vehicle but it’s not means that the vehicles are not purchased and sale in term of used vehicles. They are not ratio miner.

In Fig. 3 the category base vehicles are shows that the 0-category vehicle are mostly soled and the prices of the vehicles low and insurance risk factor are low but the vehicles are having the sales same of both category 0 and 1 and the others are category 2 and 3 are almost same sales means the prices of the vehicles are not so much difference.

The percentage pi-chart for the batter understanding shown in Fig. 4. One category is 0 category which is 32.7% and the second one is the category is 1 which is 26.3% the difference of these is 6.4% which is not the major difference but some categories are big difference like the -2 category is 1.5% this is too low symbolic sales base category. The proposed study will not to spend the major time on this category this is batter for us to take all but focus on the most demanding categories. The pi-chart helps us to what kind of category is our area of interest and what is batter for us and our proposed study approach.

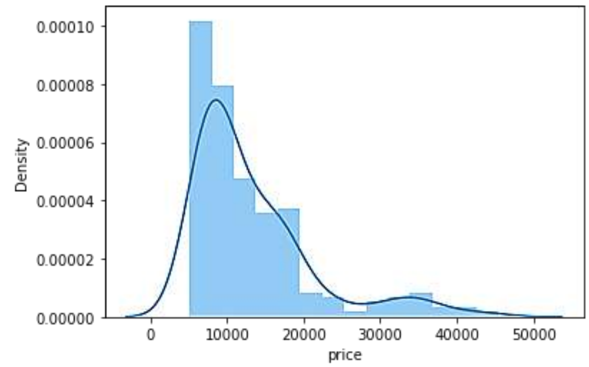

FIGURE 2: The prices base analysis of the data starts from 10,000 to 50,000.

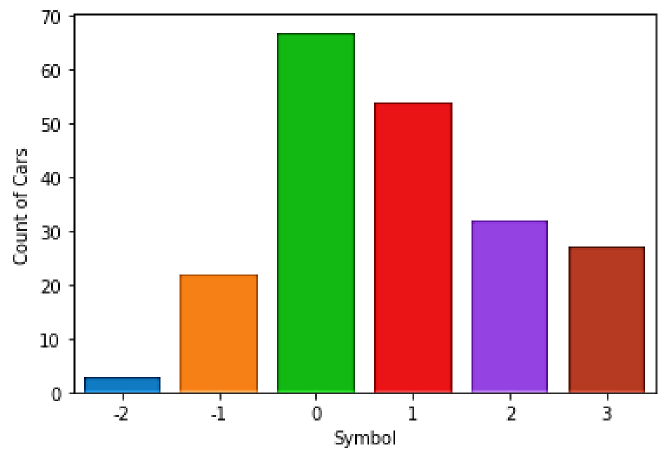

FIGURE 3: Category base sales of Vehicles the category are represented by the code from 0 to 3 and 0 to -2 each category belong to one specialized type of car category.

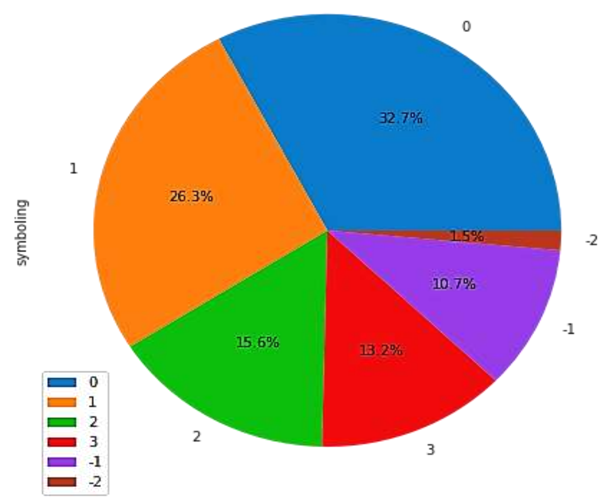

FIGURE 4: Pi-chart describe the category wise sales of the cars.

The average price of the vehicles are shown in Fig. 5 that shows the values of the vehicles are also same for the -1-Category and

3-category and the 1-category and the 2-category are almost same the category -2-category and 0-category are averagely same prices. This shows in the figure 5.0 that the values are averagely same in some but different categories in the dataset. May be the feature are miner similar or the government polices are same or the company having the same insurance and tax policies.

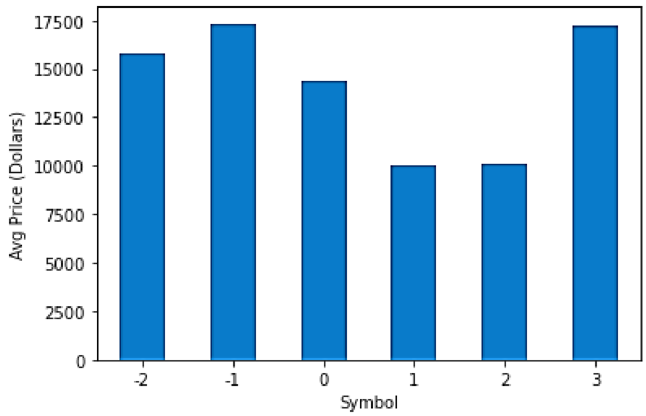

FIGURE 5.0: Category wise prices of the cars

Now need to find the models and companies of the vehicles which are consisting in the dataset and find the how much number of the vehicles are in count in the data set for the from the dataset we split the name of the company of the vehicle and model of the vehicle and then count the number of how much vehicles are in the data and the details is given in Fig. 6.

How much the type of vehicles that are using the diesel and gas and checking the how much the prices are changes from gas to diesel or fuel type if change. Figure 7 shows the histogram about the information of diesel and gas. There is a difference also $2000 averagely in the vehicles used fuel and gas the diesel type vehicles are more prices then gas. This shows that the diesel vehicles are most sales and purchased it may be related to speed pick and power and engine performance related issued but cannot ignore the gas sales which are also near to diesel sales and also having the good prices in USD.

The big one feature on that the market makes decision is milage the good milage means the market valu is good here in the figure 8.0 it shows that the the most of the vehicels are fall in the range of milage from 20-35 to are the having the mostyl sale and good sales in the market and the vehicels having the milage 15-20 are rare and laso having the expansive vehicels and the some are also cinsiderable the vehicels having the milage 45-50 that are also having the low prices and good milage but the power of the vehicels are low that’s why the people are not to buy and sales the vehicels due to the speed and horse power these vehicelsa are not most demanding besides the vechicles which are having the low milage also are slow sales due to to the high prices of the vehicles.

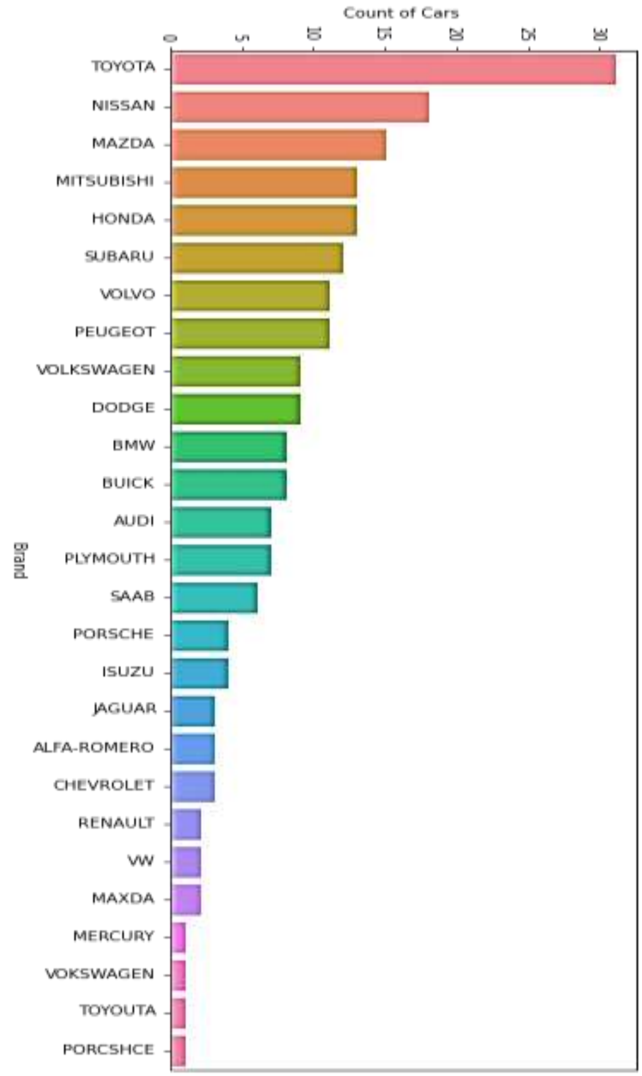

FIGURE 6.0: The data how much consist on the model wise cars and company’s wise collection.

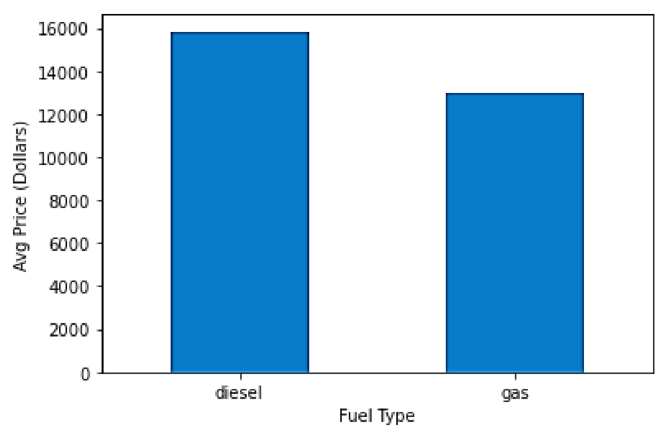

FIGURE 7: Comparison of the fuel type of the cars

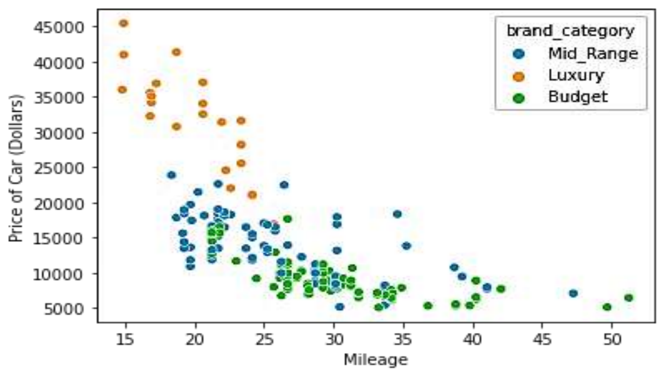

FIGURE 8: The mileage is the key stroke for decisions making for and the prices which meet the good mileage people will make decision.

Now want to check the people are what kind of vehicles type they are interested in term of fuel mileage and prices and brand that will helps us the what mind set of the people are is that the people like only under budget or they only buy expansive vehicles and they are only invested in the mid-range budgeted vehicles in Fig. 8, which shows that the peoples are not only interested in the budgeted and mid-range but the expansive vehicles are in mostly sale in the used vehicles terms people buy used vehicles only the two category one is under budget and second is mid-range they are also interested in the expansive vehicles but the so minor in the comparison of budgeted and mid-range.

V. EXPERIMETNS AND RESULTS

Here is the process model of the proposed study in Fig. 9 processes model. The processes model help us to batter understand the whole process in the pictorial format and easy to understand the processes is actually divided into data gathering and cleaning processes and second one is prepare and check the hypothesis and try the value to reach the desired results and then feature extraction and based on the feature engineering the machine learning is performed the analysis of the results are not exactly reach on the desired but after the optimal solutions are used to increase the performance of the machine leaning model.

After the univariate and bivariate analysis we are going to build the linear regression model for this purpose we use the optimal features which are we found here *FUELTYPE*, *ASPIRATION*, *CARBODY*, *DRIVEWHEEL*, *WHEELBASE*, *CARLENGTH*, *CARWIDTH*, *CURBWEIGHT*, *ENGINETYPE*, *CYLINDERNUMBER*, *ENGINESIZE*, *BORERATIO*, *HORSEPOWER*, *PRICE*, *BRAND_CATEGORY*, *MILEAGE*. Here we use label encoding for categorical attributes and encode them with 0, 1 label. The other features are rescaled for machine learning purpose the split the train and test for the machine learning here is the correlation between optimal features and the heat map shows the results in Fig. 10.

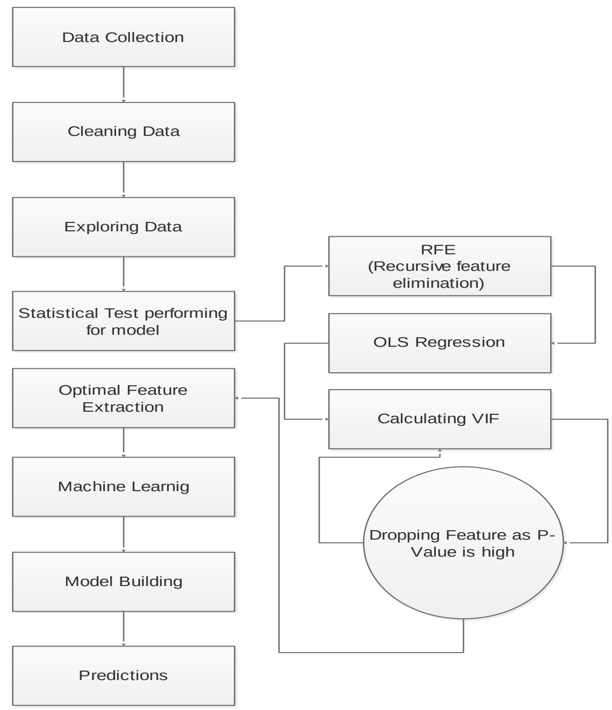

FIGURE 10: Proposed Study methodology diagram

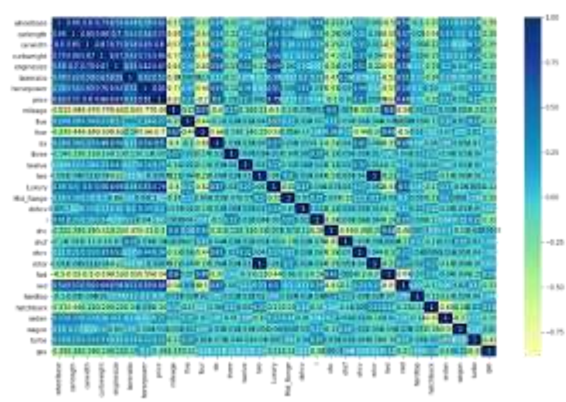

FIGURE 10: Heat map for the optimal features based.

Now have to weightage the features for this purpose we use RFE method to find the weights of the features after using the RFE for linear regression we have to find the index weighted features are *CARWIDTH*, *CURBWEIGHT*, *HORSEPOWER*, *MILAGE*, *TWELVE*, *LUXURY*, *ADHCV*, *HATCHBACK*, *SEDAN*, *WAGON*. The data is ready for machine learning model fitting after dividing the train and test we first use the statistical model in linear form for regression and find the results here the ordinary least squares method is used the method will tells us the unknown parameter in the linear regression machine learning model as shown in Fig. 11.

OLS Regression Results

Dep. Variable:	price	R-squared:	0.939
Model:	OLS	Adj. R-squared:	0.913
Method:	Least Squares	F-statistic:	162.1
Date:	Sun, 15 Jun 2021	Prob (F-statistic):	1.29e-70
Time:	07:34:45	Log-Likelihood:	205.85
No. Observations:	143	AIC:	-389.7
Df Residuals:	132	BIC:	-357.1
Df Model:	10		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-0.0947	0.042	-2.243	0.027	-0.178	-0.011
carwidth	0.2609	0.062	4.216	0.000	0.138	0.383
curbweight	0.2657	0.069	3.870	0.000	0.130	0.402
horsepower	0.4499	0.074	6.099	0.000	0.304	0.596
mileage	0.0933	0.052	1.792	0.075	-0.010	0.196
twelve	-0.1192	0.067	-1.769	0.079	-0.253	0.014
Luxury	0.2586	0.020	12.929	0.000	0.219	0.298
dohcv	-0.2676	0.079	-3.391	0.001	-0.424	-0.111
hatchback	-0.0929	0.025	-3.707	0.000	-0.143	-0.043
sedan	-0.0704	0.025	-2.833	0.005	-0.120	-0.021
wagon	-0.0997	0.028	-3.565	0.001	-0.155	-0.044

Omnibus:	43.092	Durbin-Watson:	1.867
Prob(Omnibus):	0.000	Jarque-Bera (JB):	130.642
Skew:	1.128	Prob(JB):	4.28e-29
Kurtosis:	7.103	Cond. No.	32.0

FIGURE 11: OLS-results

Standard errors accept the covariance that the covariance Metrix of the errors is accurately determined. Now are going to calculate the variance inflation factor (VIF) which gives us the multiple regression variables with the amount of multicollinearity. The value of the P is good and for this purpose we try the different techniques and dropping the mileage as p-value, dropping VIF if higher than dropping the wagon as a p value is high and still dropping the dohcv and then VIF model goes on the desired values as shown in Fig. 12.

Ordinary least squares Regression Results

Dep. Variable:	price	R-squared:	0.899
Model:	OLS	Adj. R-squared:	0.896
Method:	Least Squares	F-statistic:	308.0
Date:	Tuesday,15 Jun 2021	Prob (F-statistic):	1.04e-67
Time:	07:34:45	Log-Likelihood:	181.06
No. Observations:	143	AIC:	-352.1
Df Residuals:	138	BIC:	-337.3
Df Model:	4		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-0.0824	0.018	-4.480	0.000	-0.119	-0.046
carwidth	0.3957	0.046	8.677	0.000	0.305	0.486
horsepower	0.4402	0.052	8.390	0.000	0.336	0.544
Luxury	0.2794	0.022	12.592	0.000	0.236	0.323
hatchback	-0.0414	0.013	-3.219	0.002	-0.067	-0.016

Omnibus:	29.385	Durbin-Watson:	1.955
Prob(Omnibus):	0.000	Jarque-Bera (JB):	98.001
Skew:	0.692	Prob(JB):	5.24e-22
Kurtosis:	6.812	Cond. No.	12.9

FIGURE 12: After reducing the error rate and monitor p value for optimal values.

The error term for the analysis of residual values the normal distribution shows the error with density for the linear regression the histogram shows the distribution in Fig. 13.

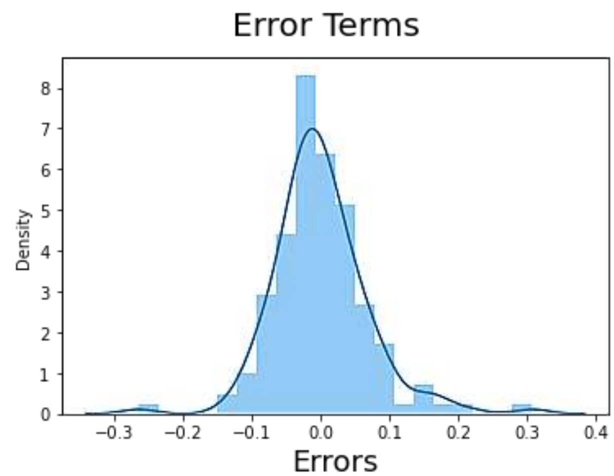

FIGURE 13: Shows the error graph is the normal distribution:

The time for prediction comes and the evaluation of the proposed model after the prediction and evaluation the results shows that the model is productive and effective. The model results the prediction with R2-Score is 90% in in the correct way and the graph in Fig. 14 shows the results.

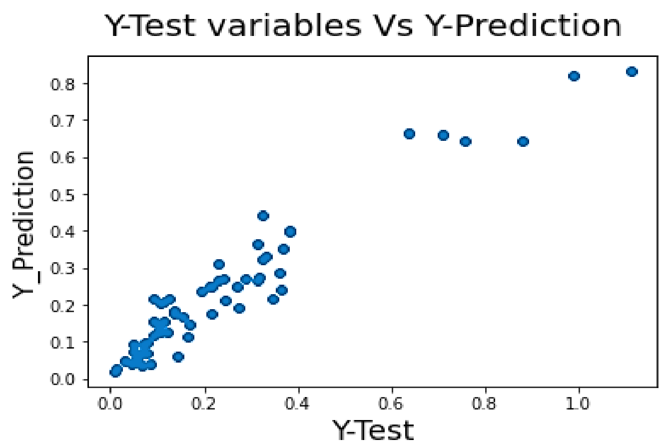

Figure 14: R^2-score is the coefficient of the determination for the regression score.

VI. CONCLUSION

The used vehicles are mostly sales in the mid-range and mostly based on the price and mileage and majorly people are not interested in the luxury vehicles but it not means the people not buy the used luxury vehicles. Proposed methodology shows the model helps the customer and dealer for customer they support them for selling and buying and as well for dealers how much price they buy the used vehicle and when they sell them the proposed system helps the dealer for good profitable deal they can make. The 90% correct prediction results show that the study is effective and efficient for both market end nodes.

REFERENCE

[1] D. Van Thai, L. N. Son, P. V. Tien, N. N. Anh, and N. T. N. Anh, "Prediction car prices using quantify qualitative data and knowledge-based system," in *2019 11th International Conference on Knowledge and Systems Engineering (KSE)*, 2019, pp. 1–5.

[2] N. Pal, P. Arora, P. Kohli, D. Sundararaman, and S. S. Palakurthy, "How much is my car worth? A methodology for predicting used cars' prices using random forest," in *Future of Information and Communication Conference*, 2018, pp. 413–422.

[3] A. Camero, J. Toutouh, D. H. Stolfi, and E. Alba, "Evolutionary deep learning for car park occupancy prediction in smart cities," in *International Conference on Learning and Intelligent Optimization*, 2018, pp. 386–401.

[4] T. Struyf *et al.*, "Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19 disease," *Cochrane Database Syst. Rev.*, Jul. 2020, doi: 10.1002/14651858.CD013665.

[5] D. Štifanić, J. Musulin, A. Miočević, S. Baressi Šegota, R. Šubić, and Z. Car, "Impact of covid-19 on forecasting stock prices: an integration of stationary wavelet transform and bidirectional long short-term memory," *Complexity*, vol. 2020, 2020.

[6] S. Sinha, R. Azim, and S. Das, "Linear Regression on Car Price Prediction," 2020.

[7] L. A. San-José, J. Sicilia, L. E. Cárdenas-Barrón, and J. M. Gutiérrez, "Optimal price and quantity under power demand pattern and non-linear holding cost," *Comput. Ind. Eng.*, vol. 129, pp. 426–434, 2019.

[8] D. Cardinale *et al.*, "Early detection of anthracycline cardiotoxicity and improvement with heart failure therapy," *Circulation*, vol. 131, no. 22, pp. 1981–1988, 2015.

[9] N. Carpentier, "'Fuck the clowns from Grease!!'Fantasies of participation and agency in the YouTube comments on a Cypriot Problem documentary," *Inf. Commun. Soc.*, vol. 17, no. 8, pp. 1001–1016, 2014.

[10] Muhammad Nasir Khan, Syed K. Hasnain, Mohsin Jamil, Sameeh Ullah, "Electronic Signals and Systems Analysis, Design and Applications International Edition," in Electronic Signals and Systems Analysis, Design and Applications: International Edition , River Publishers, 2020

[11] Khan, Muhammad Nasir, Hasnain Kashif, and Abdul Rafay. "Performance and optimization of hybrid FSO/RF communication system in varying weather." Photonic Network Communications vol. 41, no. 1, pp. 47- 56, 2021.

[12] Jamil, Mohsin, Muhammad Nasir Khan, Saqib Jamshed Rind, Qasim Awais, and Muhammad Uzair. "Neural network predictive control of vibrations in tall structure: An experimental controlled vision." Computers & Electrical Engineering, vol. 89, pp. 106940, 2021.

[13] Khan, Muhammad Nasir, Mohsin Jamil, Syed Omer Gilani, Ishtiaq Ahmad, Muhammad Uzair, and H. Omer. "Photo detector-based indoor positioning systems variants: A new look." Computers & Electrical Engineering, vol. 83, pp. 106607, 2020.

[14] Kashif, Hasnain, Muhammad Nasir Khan, and Ali Altalbe. "Hybrid optical-radio transmission system link quality: link budget analysis." IEEE Access, vol. 8, pp. 65983-65992, 2020.

[15] Khan, Muhammad Nasir, and Fawad Naseer. "IoT based university garbage monitoring system for healthy environment for students." In 2020 IEEE 14th International Conference on Semantic Computing (ICSC), pp. 354-358. IEEE, 2020.

[16] Uzair, Muhammad, R. O. B. E. R. T. D DONY, Mohsin Jamil, KHAWAJA BILAL AHMAD MAHMOOD, and Muhammad Nasir Khan. "A no-reference framework for evaluating video quality streamed through wireless network." Turkish Journal of Electrical Engineering & Computer Sciences, vol. 27, no. 5, pp. 3383-3399, 2019.

[17] Khan, Muhammad Nasir, Syed Omer Gilani, Mohsin Jamil, Abdul Rafay, Qasim Awais, Bilal A. Khawaja, Muhammad Uzair, and Abdul Waheed Malik. "Maximizing throughput of hybrid FSO-RF communication system: An algorithm." IEEE Access, vol. 6, pp. 30039-30048, 2018.

[18] Khan, Muhammad Nasir, Syed K. Hasnain, and Mohsin Jamil. Digital Signal Processing: A Breadth-first Approach. Stylus Publishing, LLC, 2016.

[19] Khan, Muhammad N. "Importance of noise models in FSO communications." EURASIP Journal on Wireless Communications and Networking vol. 2014, no. 1, pp. 1-10, 2014.

[20] M. Cocca, D. Giordano, M. Mellia, and L. Vassio, "Free floating electric car sharing: A data driven approach for system design," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 12, pp. 4691–4703, 2019.

[21] A. Barnett *et al.*, "Enacting 'more-than-human' care: Clients' and counsellors' views on the multiple affordances of chatbots in alcohol and other drug counselling," *Int. J. Drug Policy*, p. 102910, Oct. 2020, doi: 10.1016/j.drugpo.2020.102910.

[22] E. Gegic, B. Isakovic, D. Keco, Z. Masetic, and J. Kevric, "Car Price Prediction using Machine Learning Techniques," vol. 8, no. 1, p. 7, 2021.

[23] A. Fathalla, A. Salah, K. Li, K. Li, and P. Francesco, "Deep end-to-end learning for price prediction of second-hand items," *Knowl. Inf. Syst.*, vol. 62, no. 12, pp. 4541–4568, 2020.

[24] A. U. Guler, K. Misra, and V. Singh, "Heterogeneous price effects of consolidation: Evidence from the car rental industry," *Mark. Sci.*, vol. 39, no. 1, pp. 52–70, 2020.

[25] A. Pandey, V. Rastogi, and S. Singh, "Car's Selling Price Prediction using Random Forest Machine Learning Algorithm," 2020.