PRICE PREDICTION

PRAGYA SINGH PURV H JEJA VAIBHAV RATHORE

***Abstract*—A car price prediction has been a high interest research area, as it requires noticeable effort and knowledge of the field expert. Considerable number of distinct attributes are examined for the reliable and accurate prediction. To build a model for predicting the price of used cars and bikes, we applied machine learning techniques like k nearest,logical regression,linear regression . However, the mentioned techniques were applied to work as an ensemble. The data used for the prediction was collected from the kaggle.Mainly our goal is to predict the price of used cars and bikes .In this we are taking user input and according to the user input we are predicting the budget.We are asking all the inputs like what u need and which type and with the help of user input we are predicting the budget.**

I. INTRODUCTION

Predicting the price of used cars and bike in both an important and interesting problem.In many developed countries, it is common to lease a car or a bike rather than buying it outright. A lease is a binding contract between a buyer and a seller (or a third party –usually a bank, insurance firm or other financial institutions) in which the buyer must pay fixed instalments for a pre-defined number of months/years to the seller/financer. After the lease period is over, the buyer has the possibility to buy the car or bike at its residual value, When the BSVI vehicles arrived, the prices had only gone up and a few took advantage in justifying the price hikes by adding new features or refreshing the entire package. The brands will be in a position to pass on their massive decline in sales to the customers as price hikes between 2 and 5 per cent are expected over the existing sticker tags as early as next month. However, some brands could optimistically use the scenario and bring in new launches to have a first mover advantage ahead of the festive season. With the C-segment sedan sales shrinking, Honda could bring in the fifth generation City in a hope to revive its sales fortunes while Nissan's Magnite compact SUV is expected to arrive in August 2020.

The Husqvarna Svartpilen 250 and the Vitpilen 250 have become dearer for the first time by Rs. 4,736 as Bajaj Auto recently released updated prices. The similar scenario will more likely prevail across most of the two-wheeler brands by June 2020.

Accurate car and bike price prediction involves

Fig. 1. index

expert knowledge, because price usually depends on many distinctive features and factors. Typically, most significant ones are brand and model, age, kilometer and mileage. The fuel type used in the car and bike as well as fuel consumption per mile highly affect price of a car and bike due to a frequent changes in the price of a fuel. Different features like exterior color, door number, type of transmission, dimensions, safety, air condition , interior , whether it has navigation or not will also influence the car price.Different features like exterior color,number of gear, will also influence the bike price In this paper, we applied different methods and techniques in order to achieve higher precision of the used car and bike price prediction.

This paper is organized in the following manner: Section II,OUR PLAN . Section III elaborates various machine learning algorithms and examine their respective performances to predict the price of the used cars. in section IV, a conclusion of our work are given, Finally, section V, result.

II. OUR PLAN

The main aim of our project is to predict the price of used cars and bikes.In this we are taking the user input in the user input we are asking the needs like first we are asking what you need **car or bike**.

If the user clicks on car then it will ask for the model because there are several models .so, it will ask which model you want and after entering the model it will ask for the transmission like there are two types of transmission so,it will ask which transmission either manual or automatic.It will ask for how much sold vechile you want either **second hand.third hand or fourth hand** After that it will ask for the fuel as there are four types of fuel then user have to select which type user wants either petrol or diesel or LPG or CNG.then it will ask for the kilometer driven .After user enter the kilometer then finally mileage how much mileage car he/she wants .After taking all the values from the user it will predict the budget.It will also tell a how much accurate data it is.

If the user clicks on bike then it will ask for the model because there are several models .so, it will ask which model you want and after entering the model.It will ask for how much sold vechile you want either **second hand.third hand or fourth hand** After that it will ask for the kilometer driven .After user enter the kilometer then .After taking all the values from the user it will predict the budget.It will also tell a how much accurate data it is.

III. MATERIALS AND METHODS

Approach for car price prediction proposed in this paper is composed of several steps,

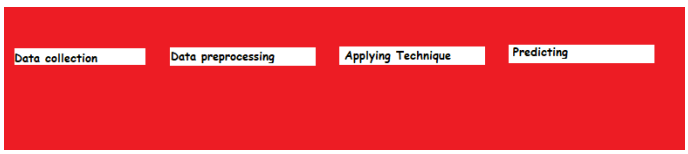

Fig. 2. Block Diagram

Data is collected from various sources like kaagle.we see which data having the prize of cars and bikes.After collectting the data we processed it.**when we collected the data for some data we didnot get the milage,then we predict the mileage from the other data .we have lots of data and half column data is missing,then we predicted the data from the rest,to make our prediction more accurate.** The following attributes were captured for each car and bike : brand, model transmission type, millage. Since manual data collection is time consuming task, especially when there are numerous records to process, a "web scraper" as a part of this research is created to get this job done automatically and reduce the time for data gathering. Web scraping is well known technique to extract information from websites and save data into local file or database. Manual data extraction is time consuming and therefore web scrapers are used to do this job in a fraction of time. Web scrapers are programed for specific websites and can mimic regular users from website's point of view After raw data has been collected and stored to local database, data preprocessing step was applied. Many of the attributes were sparse and they do not contain useful information for prediction Data are collected from

name	selling_pri	year	seller_type	owner	km_driven	ex_showroom_price
Royal Enfi	175000	2019	Individual	1st owner	350	
Honda Dio	45000	2017	Individual	1st owner	5650	
Royal Enfi	150000	2018	Individual	1st owner	12000	148114
Yamaha Fa	65000	2015	Individual	1st owner	23000	89643
Yamaha Sz	20000	2011	Individual	2nd owner	21000	
Honda CB	18000	2010	Individual	1st owner	60000	53857
Honda CB	78500	2018	Individual	1st owner	17000	87719
Royal Enfi	180000	2008	Individual	2nd owner	39000	
Hero Honc	30000	2010	Individual	1st owner	32000	
Bajaj Disc	50000	2016	Individual	1st owner	42000	60122
Yamaha FZ	35000	2015	Individual	1st owner	32000	78712
Honda Nav	28000	2016	Individual	2nd owner	10000	47255

Fig. 3. Bike graph

name	year	kilometers	fuel_Type	transmissi	owner_Ty	mileage	selling_pric
Maruti Wa	2010	72000	CNG	Manual	First	26.6 km/k	175000
Hyundai Cr	2015	41000	Diesel	Manual	First	19.67 kmp	1250000
Honda Jaz	2011	46000	Petrol	Manual	First	18.2 kmpl	450000
Maruti Erti	2012	87000	Diesel	Manual	First	20.77 kmp	600000
Audi A4 Ne	2013	40670	Diesel	Automatic	Second	15.2 kmpl	1774000
Hyundai EC	2012	75000	LPG	Manual	First	21.1 km/k	235000
Nissan Mic	2013	86999	Diesel	Manual	First	23.08 kmp	350000
Toyota Inr	2016	36000	Diesel	Automatic	First	11.36 kmp	1750000
Volkswage	2013	64430	Diesel	Manual	First	20.54 kmp	520000
Tata Indica	2012	65932	Diesel	Manual	Second	22.3 kmpl	195000
Maruti Cia	2018	25692	Petrol	Manual	First	21.56 kmp	995000
Honda City	2012	60000	Petrol	Automatic	First	16.8 kmpl	449000
Maruti Swi	2015	64424	Diesel	Manual	First	25.2 kmpl	560000

Fig. 4. Car graph

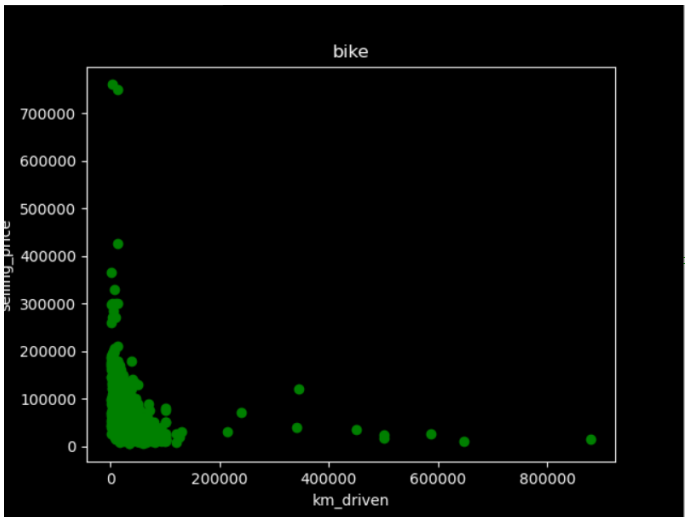

Fig. 5. Bike data

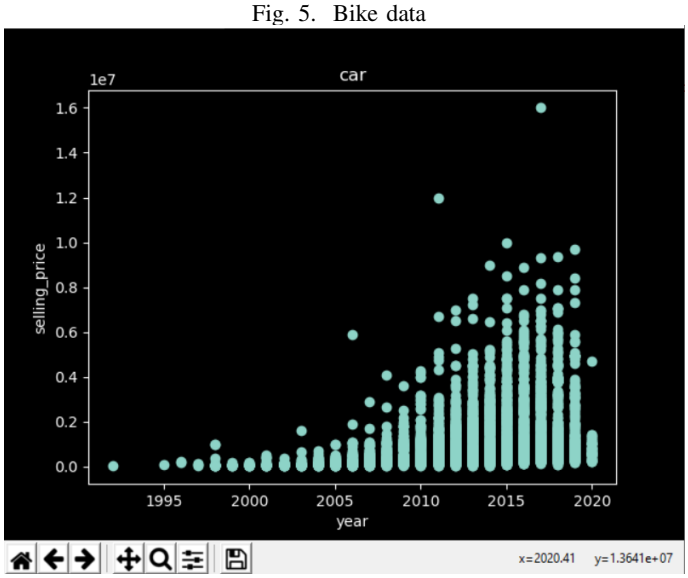

Fig. 6. Car data

kaggle or different resource.in Bike data there are 5 columns and 1000 rows and in this we also cleaned some data with the help of data cleaning.and then we applied techniques to predict the result .we used linear regression in this because data changes and the graph of this data is changing .We used sklearn modeling for train test split then we have the car data but in that half of mileage column were empty then we applied the precise algorithm so that from the other 6000 data it calculate the mileage and put in the rest of the data.After that we used data cleaning to remove the column which we do not need in the predicting price,after that we removed the kmpl from mileage column,we made a column new transmission and it will automatic put the value 1 and manual 0.We made a new fuel type like if user press 0 then it will take new fuel as LPG,if user is pressing 1 then it will take new fuel as CNG,if user is pressing 2 then it will take new fuel a PETROL and if user is pressing 4 then it will take new fuel as DIESEL.We dropped the empty data after that we used train test split and use regression to predict and with the help of linear regression w e split in train test data and we used random state as 0 and with that we are calculating the accuracy.

For the front end we used html and we connect both front end and backend by flask.

IV. CONCLUSION

Car bike price prediction can be a challenging task due to the high number of attributes that should be considered for the accurate prediction. The major step in the prediction process is collection and preprocessing of the data.Data cleaning is one of the processes that increases prediction performance, yet insufficient for the cases of complex data sets as the one in this research. Applying single machine algorithm on the data set accuracy was less than 50ensemble of multiple machine learning algorithms has been proposed and this combination of ML methods gains accuracy of 80significant improvement compared to single machine learning method approach. However, the drawback of the proposed system is that it consumes much more computational resources than single machine learning algorithm.

V. RESULT

As you can see it is difficult to think the right budget ana there a ;lot of information by which user get confused but it will help the user for making right budget for thie car or bike. Just it will take the user input and according to the user input it will predict the price of user car and bike.it will also tell how much accurate it is.