FeyNN Labs

Used Car Price Prediction

Business Model (Monetization Idea)

PROJECT 1

Harish. B

Date : 31-MAY-2023

1.0 Abstract:

The car market is highly dynamic, with prices fluctuating based on various factors such as model, brand, mileage, age, and market demand. Accurately predicting car prices is crucial for both sellers and buyers to make informed decisions and negotiate fair deals. This abstract presents a business model that utilizes Artificial Intelligence (AI) and Machine Learning (ML) techniques to predict car prices, enabling individuals and businesses to optimize their buying and selling strategies.

The proposed business model leverages a vast dataset comprising historical car sales records, incorporating features such as make, model, year, mileage, condition, and location. Using this dataset, a machine learning algorithm is trained to learn the patterns and correlations between these features and the corresponding car prices. The model employs a variety of regression techniques, such as linear regression, random forest, or gradient boosting, to achieve high prediction accuracy. To implement this business model, a user-friendly web application is developed, allowing users to input relevant car information, such as make, model, year, mileage, and condition. The AI system processes these inputs using the trained ML model and generates an estimated price range based on the learned patterns. Additionally, the system provides users with insights into the factors influencing the predicted price, such as age, mileage, and demand trends.

The Car Price Prediction business model offers several advantages. Firstly, it empowers individual car sellers by providing them with an accurate estimate of the market value of their vehicles, enabling them to set competitive prices and attract potential buyers. Secondly, it assists car buyers in making informed decisions by estimating fair prices and detecting overpriced listings. Moreover, the model can be utilized by car dealerships and automotive marketplaces to automate price estimation, optimize inventory management, and enhance customer satisfaction.

The implementation of this business model involves a scalable and adaptable infrastructure capable of handling large volumes of data. Furthermore, continuous monitoring and updates to the ML model are essential to ensure accurate predictions as market dynamics evolve over time.
In conclusion, the Car Price Prediction business model utilizing AI and ML techniques offers significant benefits to individuals and businesses operating in the car market. By providing accurate price predictions, it enhances transparency, efficiency, and fairness in car transactions, ultimately empowering buyers and sellers to make informed decisions.

I will be trying to predict the price of used cars based on their features. As it would help the people to decide whether the used car is worth the posted price by different online used - car sites. It would also help people when they plan selling their cars. When purchasing and selling, it's critical to understand the true market value of a car. A used car price prediction system is needed to accurately estimate the cars worthiness based on the features.

It's a great mistake to pay too much or sell for less than the market worth in this process so it is important to understand the price.

Introduction:

The automotive industry has witnessed tremendous growth and transformation over the years, with an increasing number of people looking to buy or sell cars. However, determining the appropriate price for a vehicle can be a complex task, as it depends on various factors such as make, model, age, mileage, condition, and market demand. Inaccurate pricing can lead to missed opportunities, financial losses, or dissatisfied customers.

To address this challenge, advancements in Artificial Intelligence (AI) and Machine Learning (ML) techniques have paved the way for more accurate and reliable car price predictions. By harnessing the power of data and algorithms, businesses can now leverage these technologies to develop robust models that estimate the value of a car based on its specific attributes.

This introduction presents a business model focused on car price prediction using AI and ML. By harnessing historical car sales data and employing sophisticated regression techniques, this model aims to provide individuals and businesses with a powerful tool to determine fair and competitive prices for their vehicles. The model not only assists sellers in setting appropriate prices but also empowers buyers to make informed decisions and identify reasonable offers in the market.

The proposed business model includes the development of a user-friendly web application that integrates with the AI system. Users can input relevant information about the car they wish to sell or purchase, such as the make, model, year, mileage, and condition. Through the trained ML model, the system generates a predicted price range, taking into account the historical data patterns and correlations. Additionally, it offers insights into the factors influencing the estimated price, enabling users to understand the valuation process better.

The implementation of this business model holds significant potential for various stakeholders in the automotive industry. Individual car sellers can benefit from accurate price estimations, attracting potential buyers and optimizing their selling strategies. Car buyers, on the other hand, can make more informed decisions by assessing whether a listing is reasonably priced or identifying potential negotiation opportunities. Additionally, car dealerships and automotive marketplaces can automate the pricing process, enhance inventory management, and improve customer satisfaction.

To ensure the effectiveness of the business model, a scalable and adaptable infrastructure is necessary to handle large datasets and accommodate future growth. Regular monitoring and updates to the ML model are also crucial to account for evolving market dynamics and maintain high prediction accuracy.

In conclusion, the Car Price Prediction business model utilizing AI and ML technologies addresses the challenges faced by both sellers and buyers in the car market. By leveraging historical data and advanced algorithms, it provides valuable insights into fair and

competitive prices, fostering transparency, efficiency, and fairness in car transactions.

The manufacturer sets the price of new cars in the industry, with the government incurring some additional costs in the form of taxes. Due to rising new car prices, used car sales are on the rise worldwide. However, because of the affordability of used cars, people tend more purchasing the used cars. When purchasing and selling, it's critical to understand the true market value of a car. A used car price prediction system is needed to accurately estimate the cars' worthiness based on the features. While there are websites that provide this service, their forecast approach may or may not be more accurate. So, several methods and algorithms may be required in the prediction of a used car's true market worth.

Predicting the resale value of a car is not a simple task. It is trite knowledge that the value of used cars depends on a number of factors. The most important ones are usually the age of the car, its make (and model), the origin of the car (the original country of the manufacturer), its mileage (the number of kilometers it has run) and its horsepower. Due to rising fuel prices, fuel economy is also of prime importance. Unfortunately, in practice, most people do not know exactly how much fuel their car consumes for each km driven. Other factors such as the type of fuel it uses, the interior style, the braking system, acceleration, the volume of its cylinders

(measured in cc), safety index, its size, number of doors, paint colour, weight of the car, consumer reviews, prestigious awards won by the car manufacturer, its physical state, whether it is a sports car, whether it has cruise control, whether it is automatic or manual transmission, whether it belonged to an individual or a company and other options such as air conditioner, sound system, power steering, cosmic wheels, GPS navigator all may influence the price as well. Some special factors which buyers attach importance in Mauritius is the local of previousowners, whether the car had been involved in serious accidents and whether it is a lady-driven car. The look and feel of the car certainly contributes a lot to the price. As we can see, the pricedepends on a large number of factors. Unfortunately, information about all these factors are notalways available and the buyer must make the decision to purchase at a certain price based on few factors only.

As a result, the model established in this study could aid online web services that determine the market worth of a used car. Many people are interested in the used automobile market because they wish to sell their car or buy a used car at some point in their lives

Problem Statement :

A. Predicting the price of used cars given the features.

Three step approach to understand the problem and the approaches used :

Step 1 : What is the problem / Task ?

-What is the price of a car given its features.

Step 2: Why does the problem need to be solved? it would help people too.

- To determine the best price by comparing the prices against different online platform used - car sites.

Step 3: How would I solve the problem ? The problem is solved using

- Machine learning techniques / models.

Market/Customer/Business Need Assessment

To be able to predict used cars market value can help both buyers and sellers.

Used car sellers (dealers): They are one of the biggest target group that can be interested in resultsof this study. If used car sellers better understand what makes a car desirable, what the important features are for a used car, then they may consider this knowledge and offer a better service.

Online pricing services: There are websites that offers an estimate value of a car. They may have a good prediction model. However, having a second model may help them to give a better prediction to their users. Therefore, the model developed in this study may help online web services that tells a used car's market value.

Individuals: There are lots of individuals who are interested in the used car market at some points in their life because they wanted to sell their car or buy a used car. In this process, it's a big corner to pay tools much or sell less then it's market value.

4.0 Target Specifications and Characterization (your customer characteristic)

Our target specifications is to predict the price of used car wherein he/she can better dell. Data extraction using web scraping of website with used car data and then proceed with data cleaning and feature selection. Also, once data cleaning is done would like to understand how various features are distributed and try normalizing them if required .

Developing a regression model to predict the price of used car based on various models like linear regression model, ridge regression model and lasso regression model and cross validating the results to find the best fit model for price prediction in the case of both test and train data. Designing a web page to demonstrate the working of best fit model and provide an easy and interactive way to the user to predict the car price based on a few questions.

5.0 External Search (online information sources/references/links)

1. Green tax- https: //www.bankbazaar.com/tax/green-tax.html

2. Predicting Used Car Prices with Machine Learning Techniques- https://towardsdatascience.com/predicting-used-car-prices-with-machinelearningtechniques-8a9d8313952

3. Used-vehicle algorithmic pricing engine method and system-

https://patents.google.com/patent/US20170300991A1/en

4. International Research Journal of Engineering and Technology (IRJET) - https://www.irjet.net/archives/V8/i4/IRJET-V8I4278.pdf

Applicable Patents (Patent of Tech/Software/Framework etc you are going to use in your Product/Service idea)

FIG. illustrates an example process for calculating a price for a vehicle, according to some embodiments.

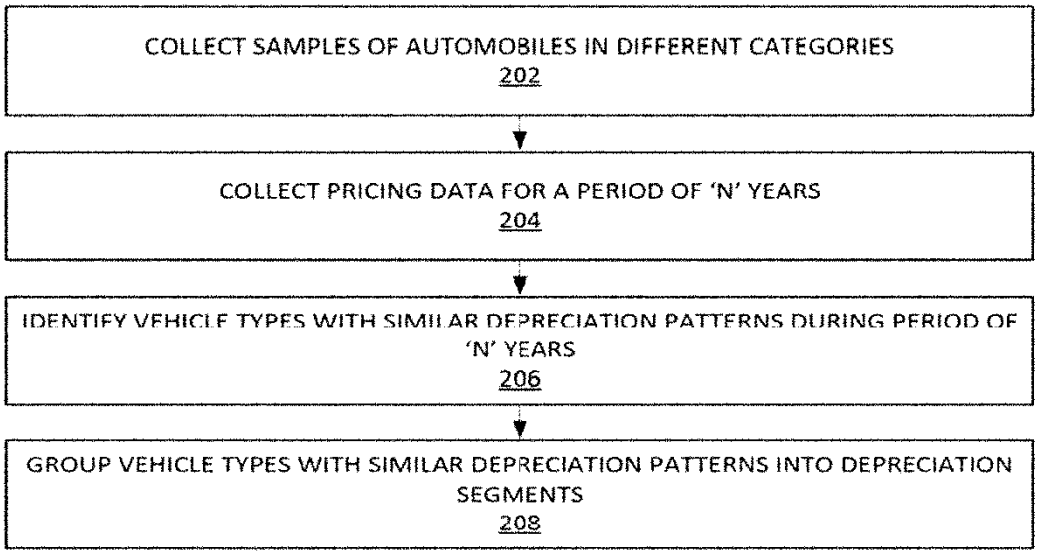
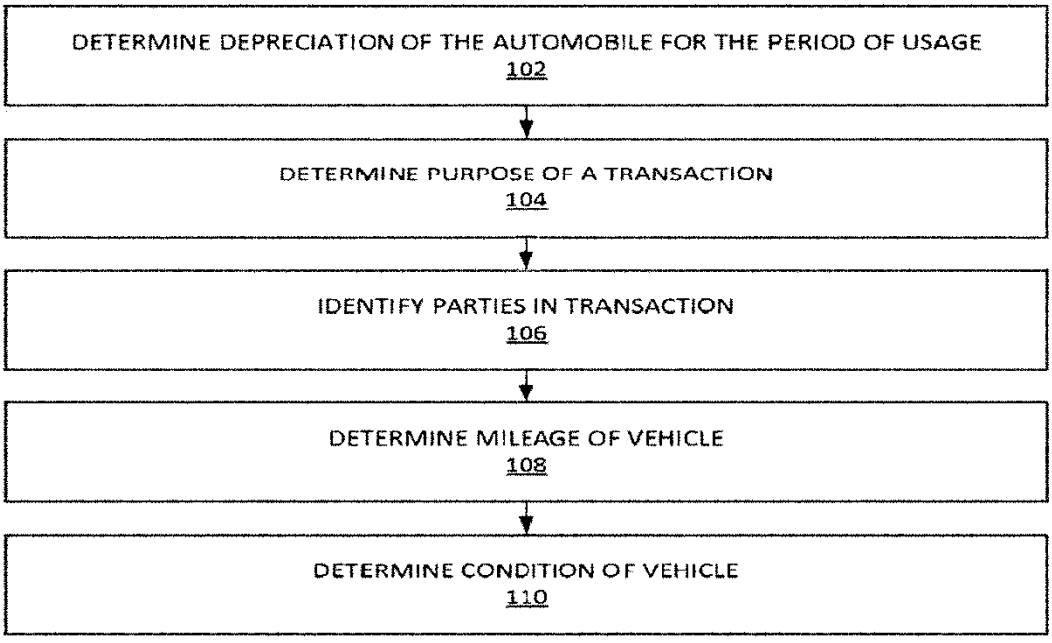


FIG. illustrates an example process of grouping vehicle types based on similar depreciation patterns, according to some embodiments.



Requirements :

Hardware requirements Operating system-

- Windows 7,8,10
- Processor- dual core 2.4 GHz (i5 or i7 series Intel processor or equivalent AMD)
- RAM-4GB

Software Requirements:

- Python
- Anoconda Navigator
- Jupyter Notebook
- Chrome
- Pip

Applicable Regulations
(government and environmental regulations imposed by countries)

Government and environmental regulations

Green Tax

Green tax is also called pollution tax or environmental tax and is the tax levied on all the goods that cause environmental pollution. It is believed that charging taxes on emissions will help bring about changes in firms and households.

As per the Latest update, the Central Government has proposed to charge a green tax on the renewal of registration certificates of personal vehicles after 15 years. The center has also proposed a lower green tax for commercial vehicles and a higher green tax for vehicles plying in highly polluted cities across the country.

Green tax for non-transport vehicles is Rs2,000 (for two-wheelers), Rs3,000 (for petrol cars) and Rs3,500 for diesel cars, while for transport vehicles, the fee is 2.5% of the yearly taxation. The tenure of green tax for non-transport vehicles is five years, while for transport vehicle is for every year.

Applicable Constraints (need for space, budget, expertise)

For building the model we need:

1. Machine learning engineer
2. Data researcher
3. Data analyst
4. Android developer

9.0 Business Model (Monetization Idea)

This model has been used by large companies, this can be extended for small businesses too. We can sell this model to seller as well. In future this machine learning model may bind with various website which can provide real time data for price prediction. Also we may add large historical data of car price which can help to improve accuracy of the machine learning model. We can build an **android app** as user interface for interacting with user. For better performance, we plan to judiciously design deep learning network structures, use adaptive learning rates and train on clusters of data rather than the whole dataset.

10.0 Concept Generation (process of coming up with Idea)

Generally, when someone wishes to purchase a used vehicle (e.g. an automobile, etc.), the user can seek the lowest price. Additionally, when selling a used vehicle, the user can seek the highest price possible. It is also a common scenario that when someone is buying a used automobile from an individual seller, the buyer can acquire the used vehicle at a much lower price than buying from an automobile dealer considering the profit margin of the dealer in the transitional transaction. Similarly, when a user is selling a used vehicle, the used vehicle can fetch a better value when the sale is made to an individual buyer than an automobile dealer as the automobile dealer would try and acquire the vehicle at a lower price and add his/her profit margin during the transitional sale. However, individual users may not have the information to maximize their quoted prices to offer their used vehicle at. Additionally, a buying non-professional user may not have sufficient information to determine a reasonable price to purchase a used vehicle.

Concept Development (Brief summary of Product/Service will be developed)

In this project we used some machine learning algorithm like Linear regression, Ridge regression, Lasso regression. All process will be conduct under python (jupyter notebook). We will use some python library like padas, numpy, seaborn.

Algorithm with highest accuracy among classification algorithms is chosen as the best algorithm for used car price prediction.

There are two primary phases in the system:

Training phase: The system is trained by using the data in the data set and fits a model (line/curve) based on the algorithm chosen accordingly.

Testing phase: the system is provided with the inputs and is tested for its working. The accuracy is checked. And therefore, the data that is used to train the model or test it, has to be appropriate.

The system is designed to detect and predict price of used car and hence appropriate algorithms must be used to do the two different tasks. Before the algorithms are selected for further use, different algorithms were compared for its accuracy. The well-suited one for the task was chosen.

For developing model we will use this:
Python : Python is a computer programming language often used to build websites and software, automate tasks, and conduct data analysis. Python is a general-purpose language, meaning it can be used to create a variety of different programs and isn't specialized for any specific problems.

Python libraries:
Pandas: pandas is a software library written for the Python programming language for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series. It is free software released under the three-clause BSD license
Syntax: Import pandas as pd

Numpy: NumPy is a library for the Python programming language, adding support forlarge, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays
Syntax: Import numpy as np

Seaborn: Seaborn is a library for making statistical graphics in Python. It is built on top of matplotlib and closely integrated with pandas data structures. This is a visualization tool usedto demonstrate the count of benign and malignant cells through the predefined dataset.
Syntax: Import seaborn as sns

Final Product Prototype (abstract) with Schematic Diagram

To develop a efficient and effective model which predicts the price of a used car according to user's inputs. To achieve good accuracy. To develop a User Interface (UI) which is user-friendly and takes input from the user and predicts the price.

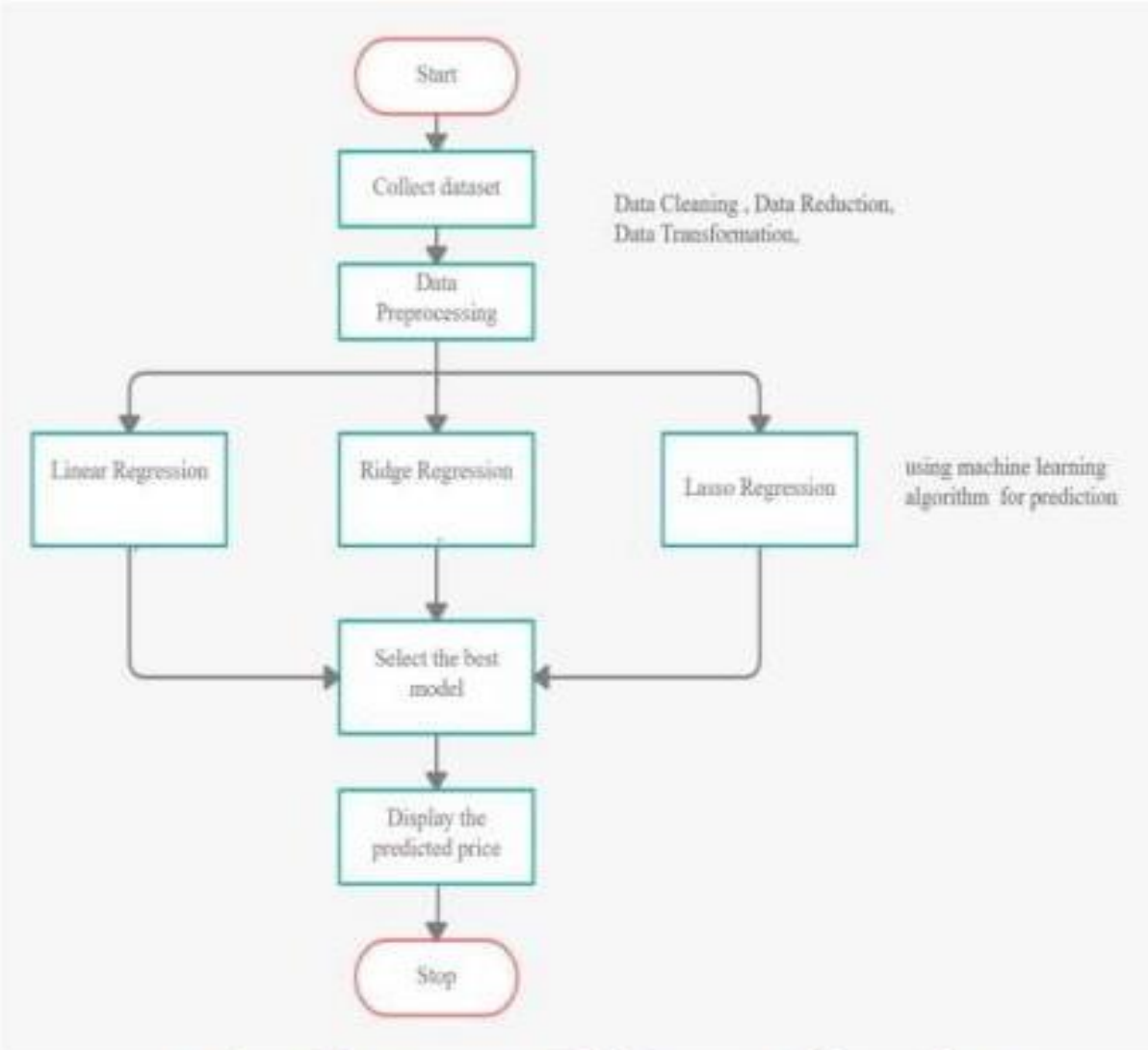


Fig – 1: Proposed System Flowchart

As shown in the above figure, the process starts by collecting the dataset. The next step is to do Data Preprocessing which includes Data cleaning, Data reduction, Data Transformation. Then, using various machine learning algorithms we will predict the price. The algorithms involve Linear Regression, Ridge Regression and Lasso Regression. The best model which predicts the most accurate price is selected. After selection of the best model the predicted price is displayed to the user according to user's inputs. User can give input through website to for used car price prediction to machine learning model.

Linear Regression

Linear Regression attempt to model the relationship between two variables by fitting a linear equation to observed data. The other is considered to be dependent variable. For Example: A modeler might want to relate weights of individuals to their heights using a linear regression model.

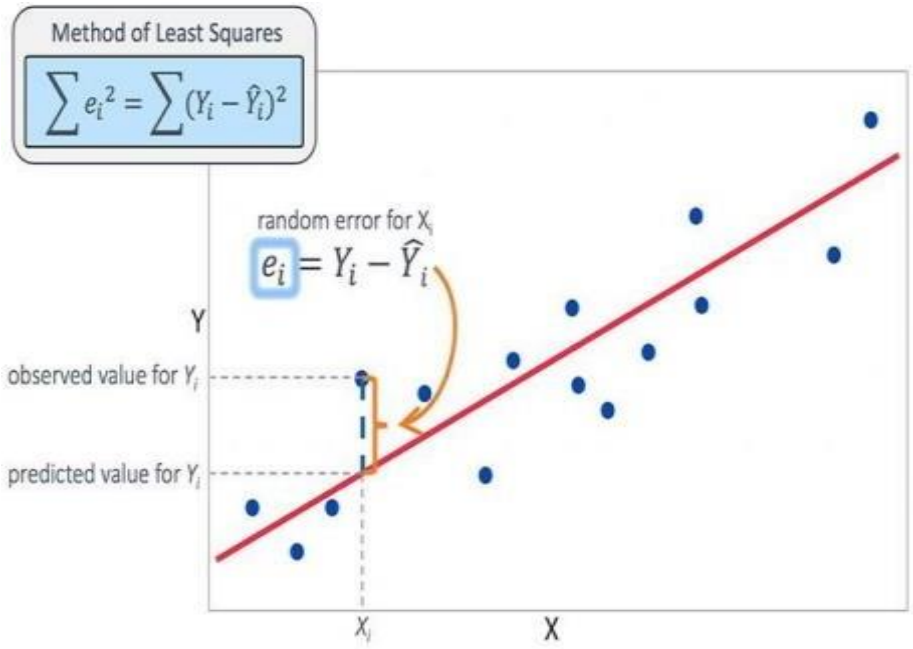


Fig – 2: Linear Regression

Linear regression is useful for finding relationship between multiple continuous variables There are multiple independent variables and single independent variable.

Y = m1X1+m2X2+……+b

m1, m2, m3 …. = Slope

b = y intercept

X1, X2, X3 …… = Independent variables

y = Dependent variables.

Ridge Regression

A Ridge regressor is basically a regularized version of Linear Regressor. The regularized term has the parameter 'alpha' which controls the regularization of the model i.e helps in reducing the variance of the estimates.

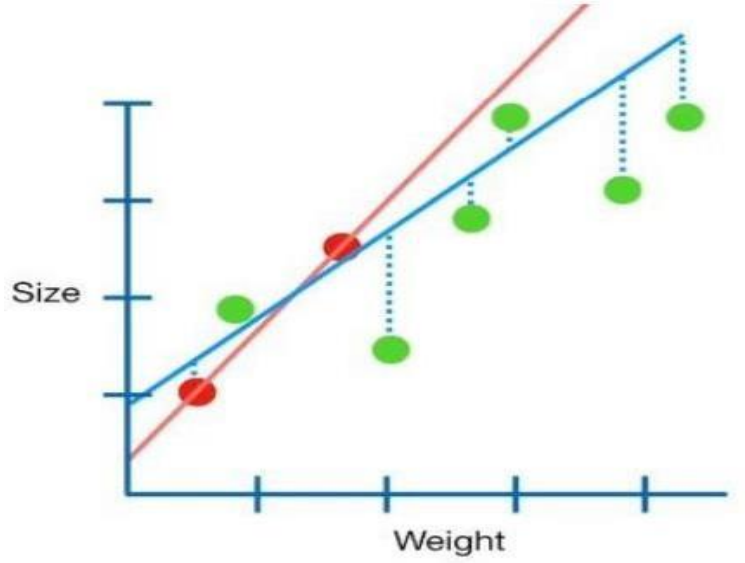


Fig – Ridge Regression

Lasso Regression

The "LASSO" stands for Least Absolute Shrinkage and Selection Operator. Lasso regression is a regularization technique. It is used over regression methods for a more accurate prediction. This model uses shrinkage. Shrinkage is where data values are shrunk towards a central point as the mean. The lasso procedure encourages simple, sparse models (i.e. models with fewer parameters). This particular type of regression is well-suited for models showing high levels of multi collinearity or when you want to automate certain parts of model selection, like variable selection/parameter elimination.

13.0 CONCLUSION

The increased prices of new cars and the financial incapability of the customers to buy them, Used Car sales are on a global increase. Therefore, there is an urgent need for a Used Car Price Prediction system which effectively determines the worthiness of the car using a variety of features. The proposed system will help to determine the accurate price of used car price prediction. This paper compares 3 different algorithms for machine learning : Linear Regression, Lasso Regression and Ridge Regression.

Feature Engineering

Feature engineering is a critical step in building an effective car price prediction model using AI and ML. It involves selecting and transforming relevant data attributes (features) to enhance the predictive power of the model. Here are some simple feature engineering techniques that can be applied in the context of car price prediction:

Age of the car: Instead of using the absolute year of manufacture, calculating the age of the car by subtracting the manufacturing year from the current year can provide a more meaningful feature. This captures the depreciation factor associated with the age of the vehicle.

Mileage normalization: Scaling or normalizing the mileage feature can help bring consistency to the data. This can be achieved by dividing the mileage by the maximum mileage in the dataset, or by using techniques such as min-max scaling or z-score normalization.

Brand and model-specific features: Extracting additional features related to the brand and model of the car can provide valuable information. For example, creating binary variables indicating the presence of popular features or luxury attributes specific to certain brands or models can capture the influence of brand reputation on pricing.

Historical price trends: Incorporating historical price trends can capture market dynamics and price fluctuations over time. Calculating features such as average price in the past year or the rate of price change can provide insights into how the market value of a car has evolved.

Location-based features: Including location-based features such as city or region can help account for regional price variations or demand-supply dynamics. Certain areas may have higher demand or specific preferences for certain types of cars, which can affect their prices.

Condition-based features: Transforming the condition attribute into a numerical scale, such as a rating or a quality score, can make it more amenable to the ML model. This can be done by converting categorical condition labels (e.g., excellent, good, fair) into numerical values.

Engine and performance attributes: Extracting relevant information from engine specifications, such as horsepower, torque, or fuel efficiency, can capture the influence of performance characteristics on pricing. These attributes can be transformed or aggregated to create more meaningful features.

Market demand indicators: Incorporating external factors that indicate market demand, such as search volume or online listings for a particular car model, can provide insights into its popularity and potential impact on pricing.

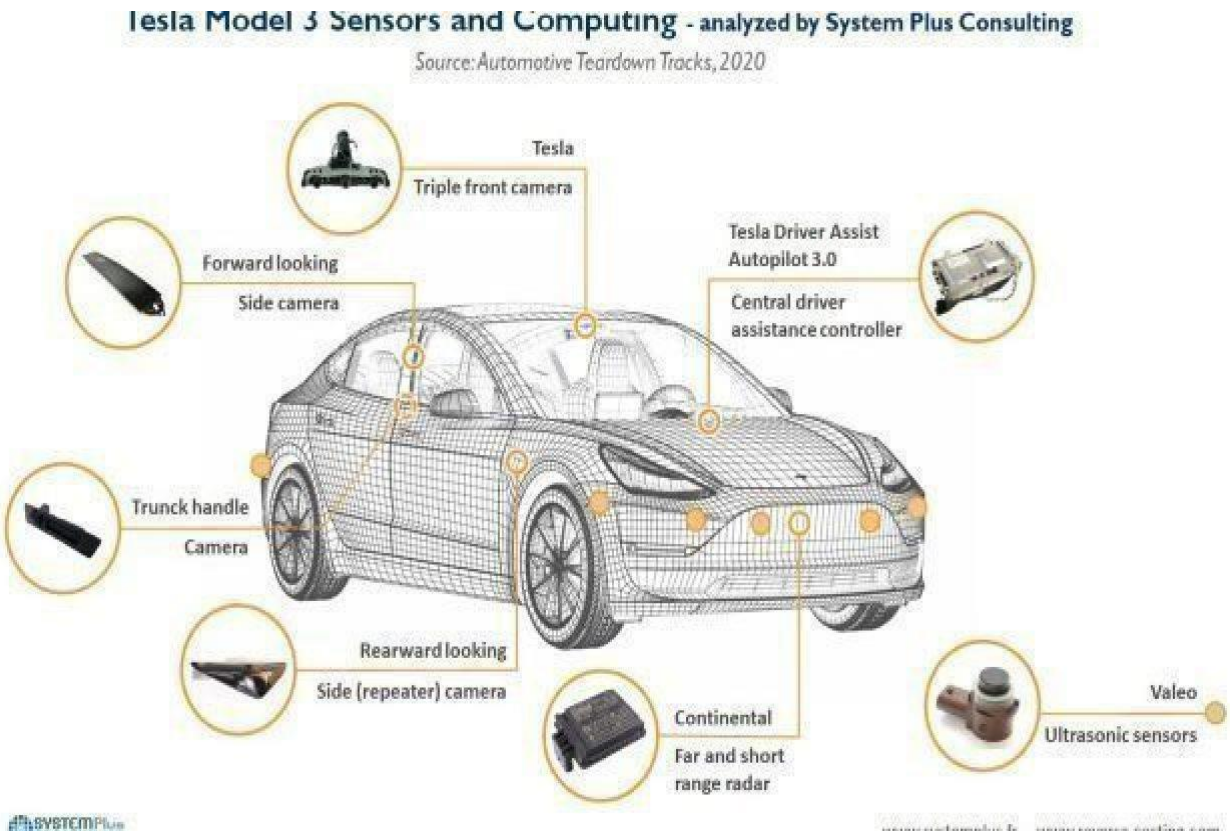
Remember, the choice of feature engineering techniques may vary depending on the specific dataset and the ML algorithms used. Exploratory data analysis and domain knowledge can guide the selection and transformation of features to optimize the model's predictive accuracy.

How Tesla uses AI and CV

While the memes and tweets kept flowing about how the popularly acclaimed "Self-driving Cars" would function in India, AI specialists in India rejoiced. This is just the beginning of an entire wave of Computer Vision taking over the world!

Most of us data science enthusiasts must be aware of some basic Computer Vision concepts like Object Detection, Tracking, etc. However, we all know reality is so much farther from mere theoretical concepts. You must be wondering about the technological stack Tesla uses in delivering what will be a fully automated car of the future! Sounds fascinating, doesn't it? (Although motor heads may not be as excited with this futuristic advancement). In this article, we discuss some high-level concepts that will make Tesla cars attain full autonomy!

REFERENCE: https://www.fireblazeaischool.in/blogs/how-tesla-uses-ai-and-cv



14.0 REFERENCES

[1] Sameerchand Pudaruth, “Predicting the Price of Used Cars using Machine Learning Techniques”;(IJICT 2014)

[2] Enis gegic, Becir Isakovic, Dino Keco, Zerina Masetic, Jasmin Kevric, ”Car Price Prediction Using Machine Learning”; (TEM Journal 2019)

[3] Ning sun, Hongxi Bai, Yuxia Geng, Huizhu Shi, “Price Evaluation Model In Second Hand Car System Based On BP Neural Network Theory”; (Hohai University Changzhou, China)

[4] Nitis Monburinon, Prajak Chertchom, Thongchai Kaewkiriya, Suwat Rungpheung, Sabir Buya, Pitchayakit Boonpou, “Prediction of Prices for Used Car by using Regression Models” (ICBIR 2018)

[5] Doan Van Thai, Luong Ngoc Son, Pham Vu Tien, Nguyen Nhat Anh, Nguyen Thi Ngoc Anh, “Prediction car prices using qualify qualitative data and knowledge-based system” (Hanoi National University)