**MIDTERM PROJECT**

**Used Cars Price Prediction**

**Team Data Wizards**

**Raparla Sravani(016656601)**

**Chirudeep (016682627)**

**Sai pragna kancheti(016698552)**

**Objective:**

· To create a supervised machine learning model for predicting a vehicle's worth based on numerous factors.

· The system being developed must be feature-based, which means that feature-wise prediction must be achievable.

· Providing graphical comparisons for a clearer picture.

To create an efficient and effective model that estimates the price of a used automobile based on the inputs of the user. To attain high precision. To create a user-friendly User Interface (UI) that takes user input and forecasts the pricing.

<https://medium.datadriveninvestor.com/end-to-end-project-on-used-car-price-prediction-3dc412d24aa0>

<https://www.kaggle.com/nehalbirla/vehicle-dataset-from-cardekho>

In the system, there are 2 major phases:

1. Training phase: Using the data in the data set, the system is trained to fit a model (line/curve) depending on the algorithm used.

2. Testing phase: the system is given inputs and its functionality is tested.

The accuracy is verified. As a result, the data that is used to train or test the model must be appropriate. Because the system is meant to identify and estimate the price of used automobiles, separate algorithms must be utilized to accomplish the two goals.

Different algorithms were assessed for their accuracy before being chosen for future usage.The well-suited one for the task was chosen.

Refinements planned to perform : Data cleaning,Data imputation,Data Preprocessing, Feature Engineering

**Models:**

* clustering(K-Means)
* Regression (Linear regression,Gradient boost)
* classification(Decision Tree, Random forest , naive bayes)

**Data analysis:**

Data analysis is a process of inspecting , cleansing, transforming, and modeling data with the goal of discovering useful information, informing conclusions, and supporting decision-making.

**Website Reference:**

**Gaadi.com**

CarDekhoGaadi entered the Indian used vehicle market to enhance the process of selling a car for you. From determining the best resale price for a used automobile to conducting loan settlement and RC transfer, They've got you covered. This way, you can sit back and relax as we utilize a combination of cutting-edge technology and seasoned personnel to find you the greatest deal on your vehicle.

**Problem Identification & objectives**

· To create a supervised machine learning model for predicting a vehicle's worth based on numerous factors.

· The system being developed must be feature-based, which means that feature-wise prediction must be achievable.

· Providing graphical comparisons for a clearer picture.

To create an efficient and effective model that estimates the price of a used automobile based on the inputs of the user. To attain high precision. To create a user-friendly User Interface (UI) that takes user input and forecasts the pricing.

**Data Visualization :**

**T**his is the most crucial phase in the data science life cycle; here, we strive to understand the behavior of data and get useful insights from it. Let's see if we can figure it out by doing.

Feature Importance refers to techniques that calculate a score for all the input features for a given model- the scores simply represent the “importance” of each feature. A higher score means that the specific feature will have a larger effect on the model that is being used to predict a certain value.