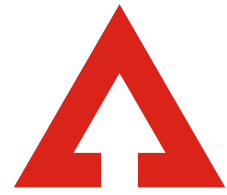# USED CAR PRICE PREDICTION

GUIDED BY

Anay Tamhankar

Trupti Joshi

Prasad Deshmukh

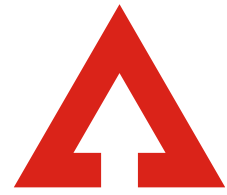# GROUP MEMBERS

Abhishek Patel   (230943025001)   Ajinkya Gaiki   (230943025003)
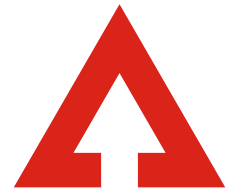
Nikhil Gaikwad   (230943025017)   Vaibhav Gurav   (230943025021)

INTRODUCTION

▶ Analyzing market trends

▶ Need for Price Prediction model that effectively determines the worthiness of the car using a variety of dependent features.

▶ Apache PySpark for data preprocessing and ML ,

▶ Tableau for Data Visualization

PROBLEM STATEMENT

▶ Lack of reliable methods to determine fair prices for cars.

▶ 50% of people prefer to buy used cars

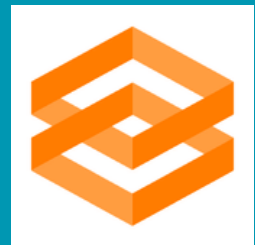▶ To predict the car price using PySpark ML model
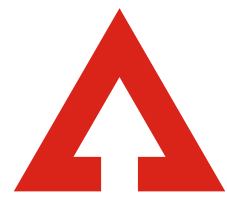
ARCHITECTURE

Data collection → Data pre-processing → Exploratory data analysis

↓

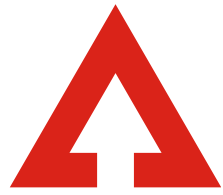Predictions & Deployment ← Model Training ← Clean Data → Data Visualization

DATA PRE-PROCESSING

1. Dropping Independent Columns:
2. Removing Duplicate Records:
3. Handling Missing Values:
4. Handling Categorical Columns:
5. Converting Data Types:
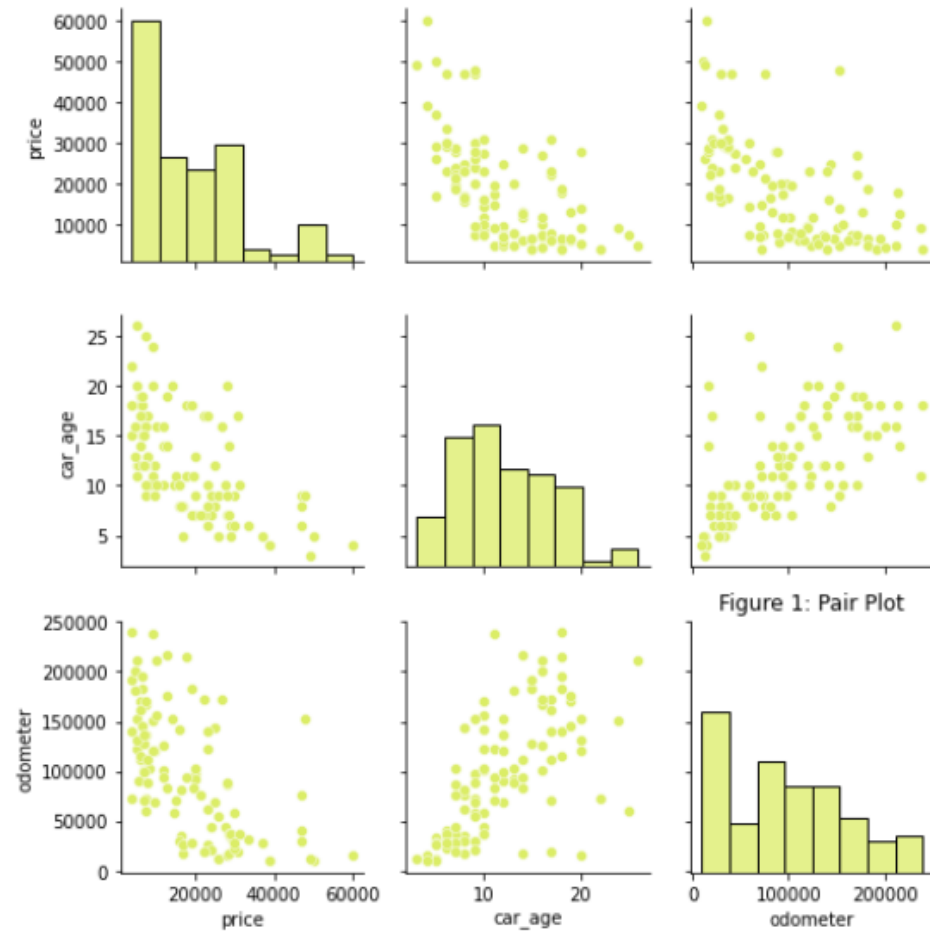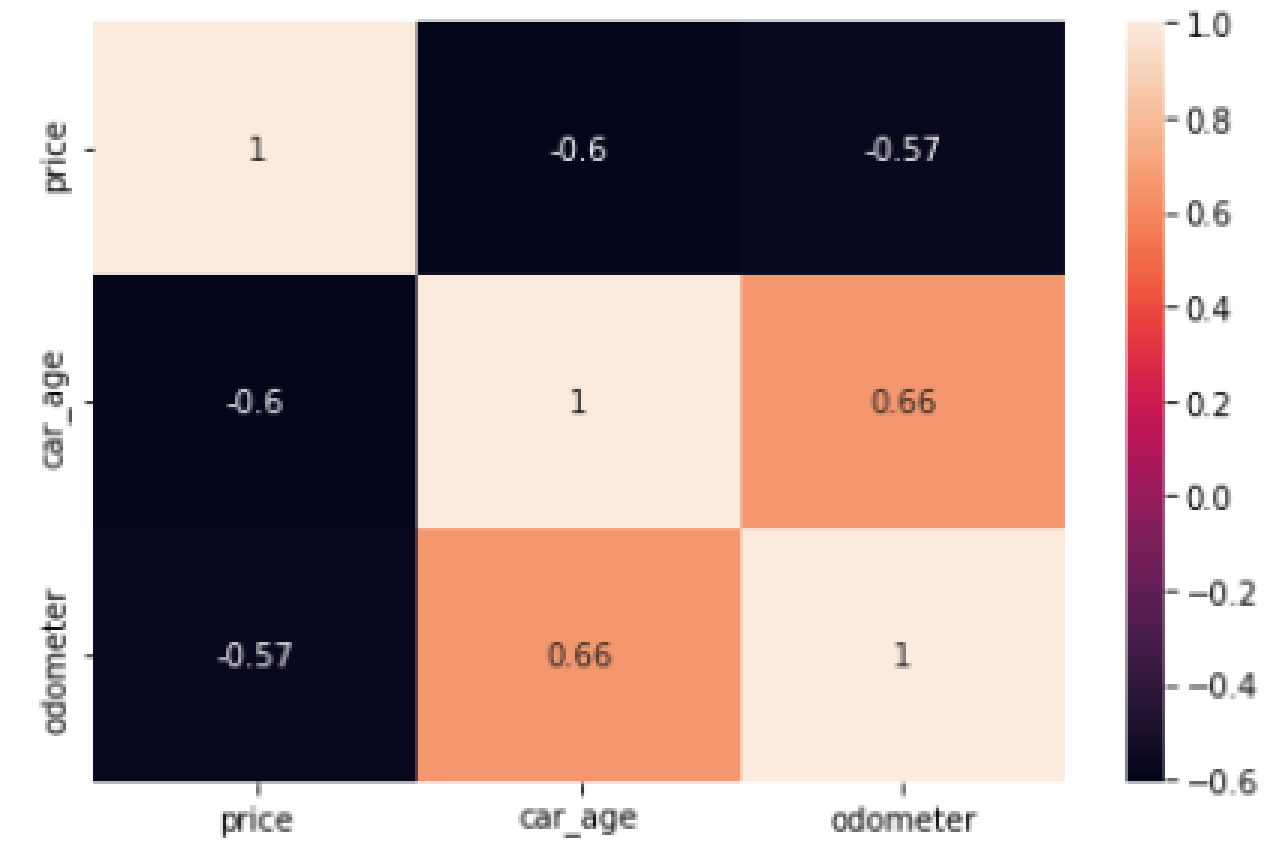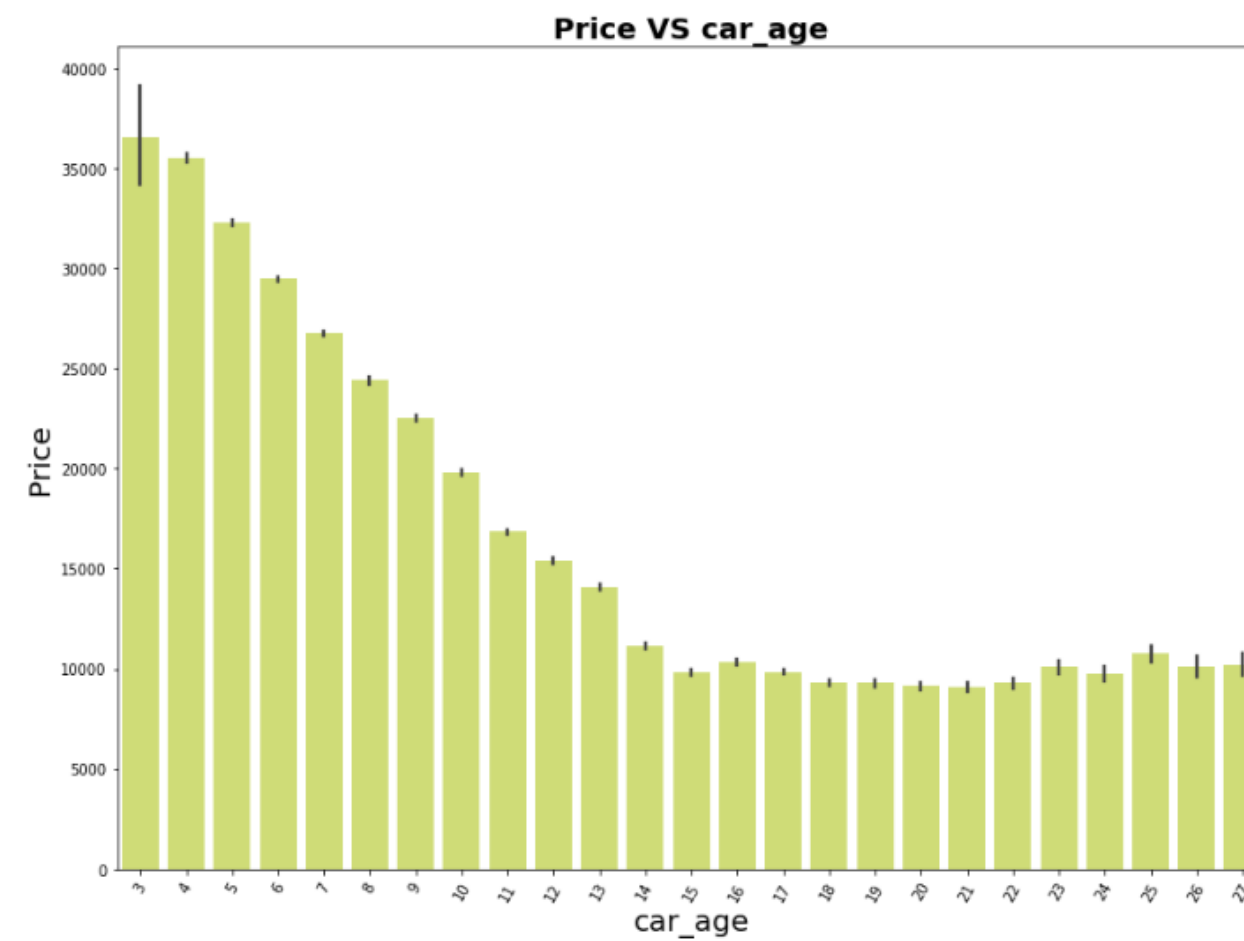6. Handling Outliers:
7. Adding New Column

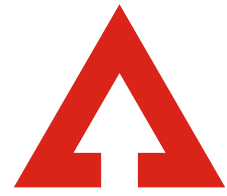# EDA

Figure 1: Pair Plot

Price VS car_age

ML MODELS

| Model	R2 Score
1. Linear Regression	0.74
2. Random Forest Regression	0.68
3. Decision Tree	0.65
4. Ridge	0.74
5. Lasso	0.74
6. Gradient Boosting	0.76
7. Gradient Boosting(T)	0.84

CONCLUSION

▶ Use of multiple Machine-learning Algorithms to build up automobile price forecasting model.

▶ Gradient Boosting is the best model demonstrates the highest prediction performance with R2 score of 0.84.

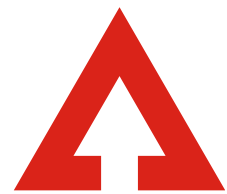▶ Provides valuable insights for both sellers and buyers .

PROBLEM FACED

▶ Pre-processing of Data

▶ Tuning of Model

▶ User Interface using gradio
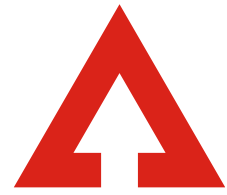
FUTURE SCOPE

▶ Can provide this model to bind with different website

▶ We can build an android app for interacting with user.

▶ For better performance, we plan to design deep learning network structures, use adaptive learning rates

THANK YOU