

Car Price Prediction Project Report

1. Project Objective

To build a machine learning model that accurately predicts the resale price of cars based on factors like original price, kms driven, transmission type, fuel type, number of owners, and car age.

2. Dataset Overview

- Total Records: 301
- Key Features: Present_Price, Kms_Driven, Owner, Fuel_Type, Selling_type, Transmission, Car_Age
- Target Variable: Selling_Price

3. Data Preprocessing Steps

- Dropped Car_Name (not useful for modeling)
- Converted Year to Car_Age (2020 - Year)
- One-hot encoded categorical variables (Fuel_Type, Selling_type, Transmission)
- Scaled and split data into training and test sets (80/20)

4. Model Training and Evaluation

Model Summary:

- Model: Linear Regression
- Features Used:
 - Present Price
 - Kms Driven
 - Owner Count

Car Price Prediction Project Report

- Fuel Type (One-Hot Encoded)
- Selling Type (One-Hot Encoded)
- Transmission Type (One-Hot Encoded)
- Car Age (2020 - Year)

Performance Metrics:

- R^2 Score: 0.83
- RMSE: 1.12 Lakhs

Visualizations Summary:

- Distribution plot of Actual vs Predicted prices shows close match
- Residual plot indicates homoscedasticity (good fit)
- Correlation heatmap confirms high impact of Present Price and Car Age on Selling Price

5. Conclusion

The linear regression model delivered strong performance with an R^2 score of 0.83. Visual diagnostics support that the model generalizes well. For further improvement, ensemble models like Random Forest or Gradient Boosted Trees can be considered. This project demonstrates the power of feature engineering and simple models in predicting car prices accurately.