Car Price Prediction Documentation

Introduction

This project predicts the selling price of a car based on features such as fuel type, transmission, present price, and car age using Linear Regression and Lasso Regression.

Project Overview

• Predicts car resale value using machine learning

• Uses Linear Regression and Lasso Regression

• Lasso provides better regularization and stability

Libraries Used

• NumPy

• Pandas

• Matplotlib and Seaborn

• Scikit Learn

Dataset and Features

The dataset contains columns like Fuel Type, Transmission, Seller Type, Present Price, Year.

Car Age is created as 2025 minus Year. Car_Name is removed.

Preprocessing

• Handling missing values and duplicates

• Encoding categorical variables

• Feature selection and Car Age creation

EDA Insights

• Older cars have reduced value

• Diesel and automatic cars generally priced higher

• Present Price strongly influences Selling Price

Models Used

Linear Regression

A basic model that provides a baseline.

Lasso Regression

Improves model by applying L1 regularization and reducing overfitting.

Evaluation Metrics

- R2 Score
- MAE
- MSE
- RMSE

Results

Lasso Regression performs better than Linear Regression in accuracy and generalization.

Conclusion

Car price prediction can be effectively modeled using Lasso Regression due to its strong feature control and reduced overfitting.

Author

Satyam Gajjar