

Car Price Prediction Using Regression Models

This project aims to predict the selling price of used cars based on various features like car year, mileage, fuel type, transmission, and more. The dataset contains car details, including features such as car name, year, selling price, present price, kilometers driven, fuel type, seller type, transmission type, and the number of previous owners.

Steps Involved:

1. Data Preprocessing:

- **Handling Missing Values:** Missing or incomplete data is handled using appropriate imputation methods.
- **Feature Encoding:** Categorical features (like fuel type, transmission, etc.) are transformed into numerical values using techniques like one-hot encoding.
- **Feature Scaling:** Features are standardized to ensure that the machine learning models do not give undue weight to any particular feature.

2. Model Selection:

- **Linear Regression:** A basic approach where we assume a linear relationship between the features and the target variable (selling price).
- **Ridge Regression:** A regularized form of linear regression to handle multicollinearity and overfitting by adding a penalty to the model.
- **Lasso Regression:** Another form of regularized regression that performs both regularization and feature selection by shrinking some coefficients to zero.

3. Model Evaluation:

- The models are evaluated using metrics such as **Mean Absolute Error (MAE)**, **Root Mean Squared Error (RMSE)**, and **R-squared** to measure the performance and accuracy of predictions.

4. Result Interpretation:

- The results are interpreted to identify which features most influence the selling price of cars, and how different regression models compare in terms of accuracy and reliability.

5. Optimization:

- Hyperparameter tuning is performed to optimize model performance. Techniques like cross-validation are used to avoid overfitting and ensure generalizability to unseen data.

Key Findings:

- The project demonstrates the effectiveness of using various regression models to predict car prices accurately.
- Regularization methods like **Ridge** and **Lasso Regression** help prevent overfitting and improve model performance.

Conclusion:

This project illustrates how machine learning can be applied to real-world business problems like pricing predictions in the automotive market. By using advanced regression models, we can develop more accurate and reliable tools for price prediction, which can be valuable for both car sellers and buyers.