**Title**: Car Price Prediction Using Machine Learning

Submitted by: GareshKumar (69677)

Course: Machine Learning with Python

Instructor: [Sulleman Ahmed Naz]

Date: [20/06/2025]

**Descrption of the Project**

This project aims to predict the selling price of used cars based on various features such as:

Year of manufacture

Present price

Kilometers driven

Fuel type (Petrol, Diesel, CNG)

Seller type (Dealer, Individual)

Transmission type (Manual, Automatic)

Number of previous owners

The dataset used contains information about different cars, including their specifications and selling prices. The goal is to build a regression model that accurately predicts the selling price of a car given its features.

3. **Methodology**

Steps Followed:

Data Collection:

The dataset (car data.csv) was loaded into a Pandas DataFrame.

Data Preprocessing:

Checked for missing values (none found).

Encoded categorical variables (Fuel\_Type, Seller\_Type, Transmission) into numerical values.

Feature Selection:

Dropped irrelevant columns (Car\_Name).

Separated features (X) and target variable (Y = Selling\_Price).

Train-Test Split:

Split the data into training (90%) and testing (10%) sets.

Model Training:

Linear Regression: A basic regression model to predict continuous values.

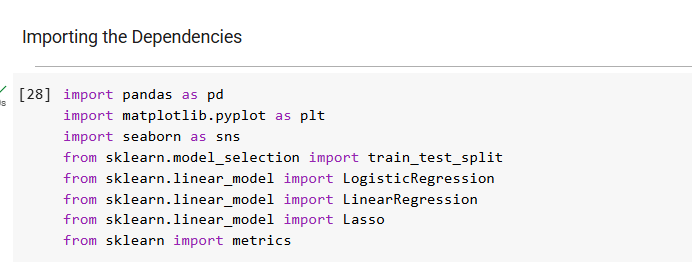
Lasso Regression: A regularized version of linear regression to prevent overfitting.

Model Evaluation:

Used R-squared error to measure model performance.

Visualized actual vs. predicted prices using scatter plots.

**4) Code**

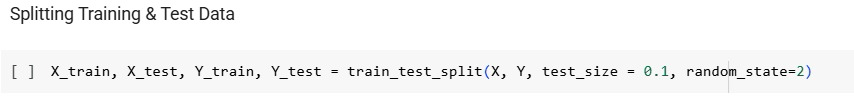
**Importing Libraries**

**Data Preprocessing**

****

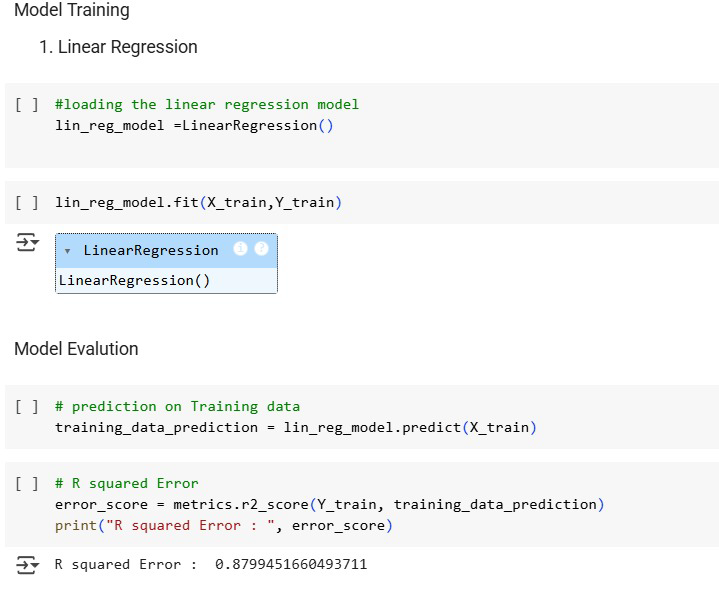
**Splitting Data**

****

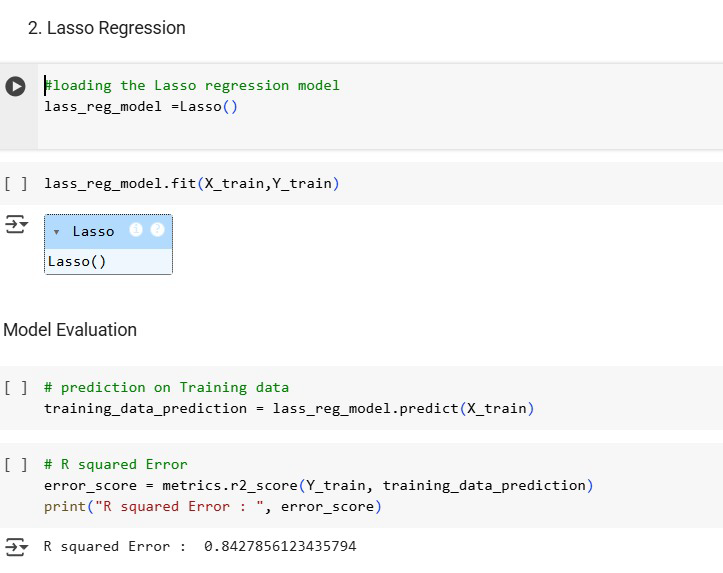
****

**Model Training & Evalutation**

**Linear Regression**

****

**Lasso Regression**

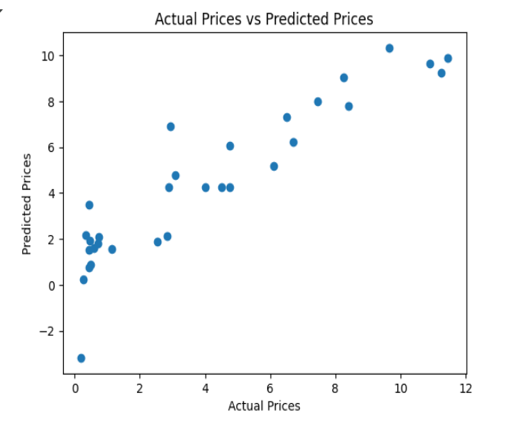
****

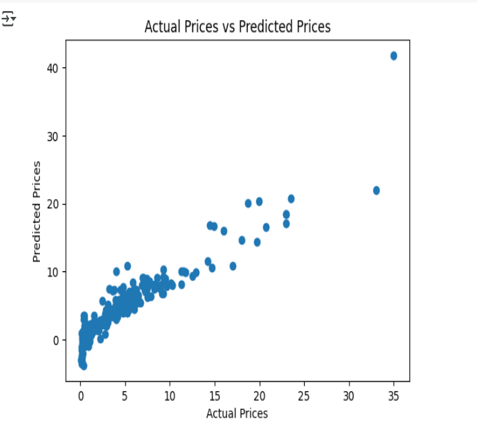
**Results**

**Linear Regression Results**:

Training R² Score: [0.8799451660493711]

Testing R² Score: [0.8365766715027051]

Scatter Plot:

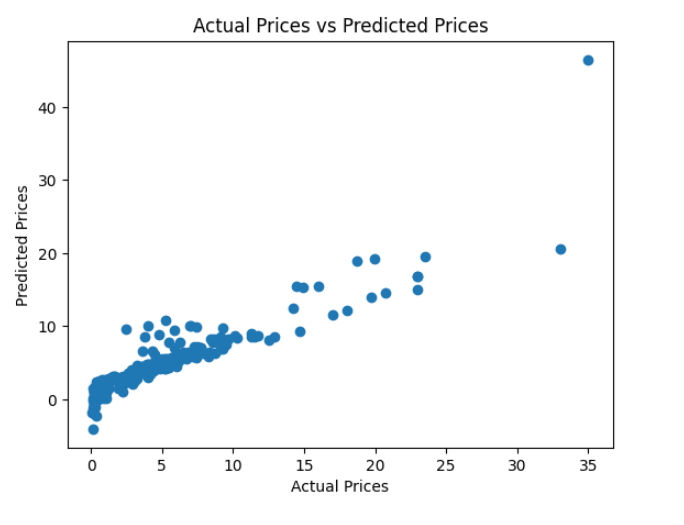
****

**Lasso Regression Results**

Training R² Score: [0.8427856123435794]

Testing R² Score: [0.8709167941173195]

Scatter Plot:



**Conclusion**

Both Linear Regression and Lasso Regression performed well, with high R² scores.

Lasso Regression may generalize better due to regularization.

Future improvements could include:

Trying other models (Random Forest, Gradient Boosting).

Feature engineering (e.g., age of car instead of year).