Graduation Project DEPI

Predicting Used Car Prices with ML

Accurate prediction of used car prices is crucial for both buyers and sellers to ensure fair transactions. This presentation will showcase a machine learning model developed to tackle this challenge, highlighting its exceptional performance in the Egyptian market.

Our Team

- **Abdulrahman Nabil**
- **Kareem Saad**
- **Abdulrahman Tarek**
- **Zyad Nabil**

Steps of Project

1. Libraries & Toolkits
2. Importing Dataset & Exploration
3. Exploratory Data Analysis (EDA)
4. Data Preprocessing & Feature Engineering
5. Feature Selection
6. Data Modeling
7. Model Evaluation & Saving
8. Deploy on Stream lit

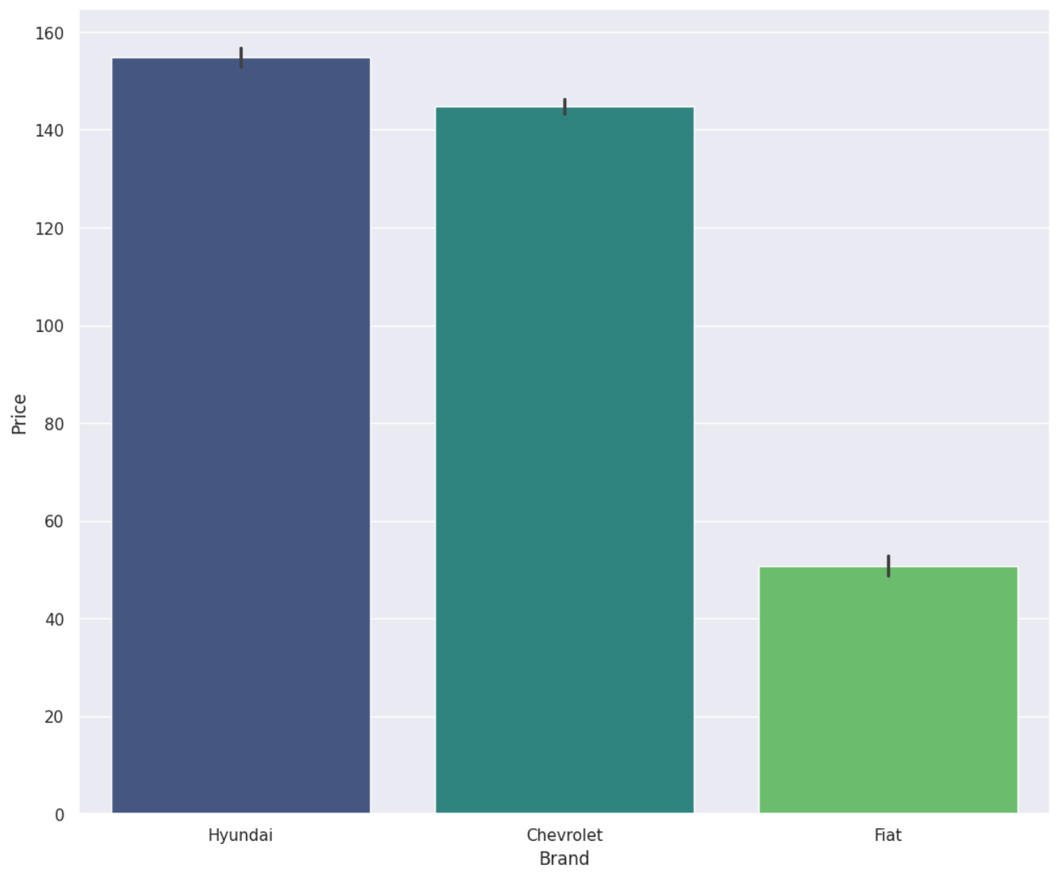
Dataset

Sample of Data

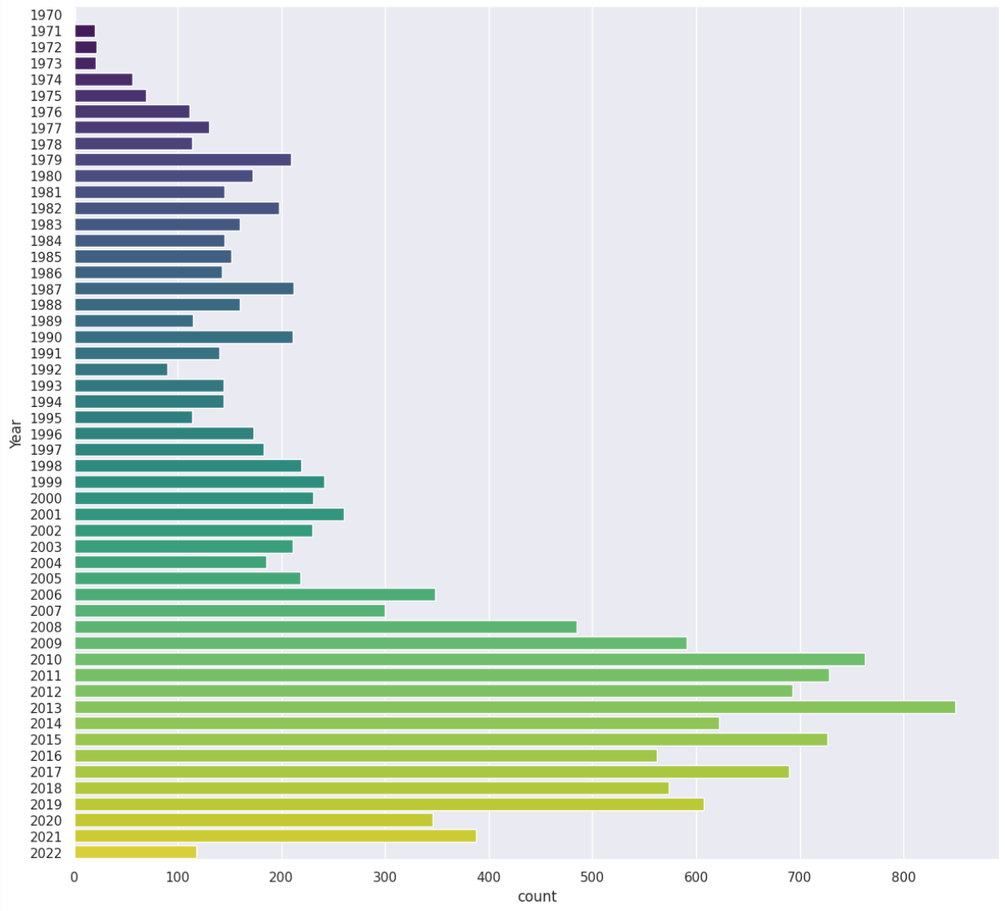
	Brand	Model	Body	Color	Year	Fuel	Kilometers	Engine	Transmission	Price	Gov
5337	Hyundai	Accent	Sedan	Black	2007	Benzine	140000 to 159999	1600 CC	Automatic	140	Giza
5338	Hyundai	Accent	Sedan	Silver	2005	Benzine	180000 to 199999	1000 - 1300 CC	Manual	78	Qena
5339	Hyundai	Accent	Sedan	Gray	1999	Benzine	140000 to 159999	1400 - 1500 CC	Manual	70	Giza
5340	Hyundai	Accent	Sedan	Blue- Navy Blue	2009	Benzine	140000 to 159999	1600 CC	Automatic	150	Cairo

Exploratory Data Analysis (EDA)

Analyze categorical variables and their relationship with price

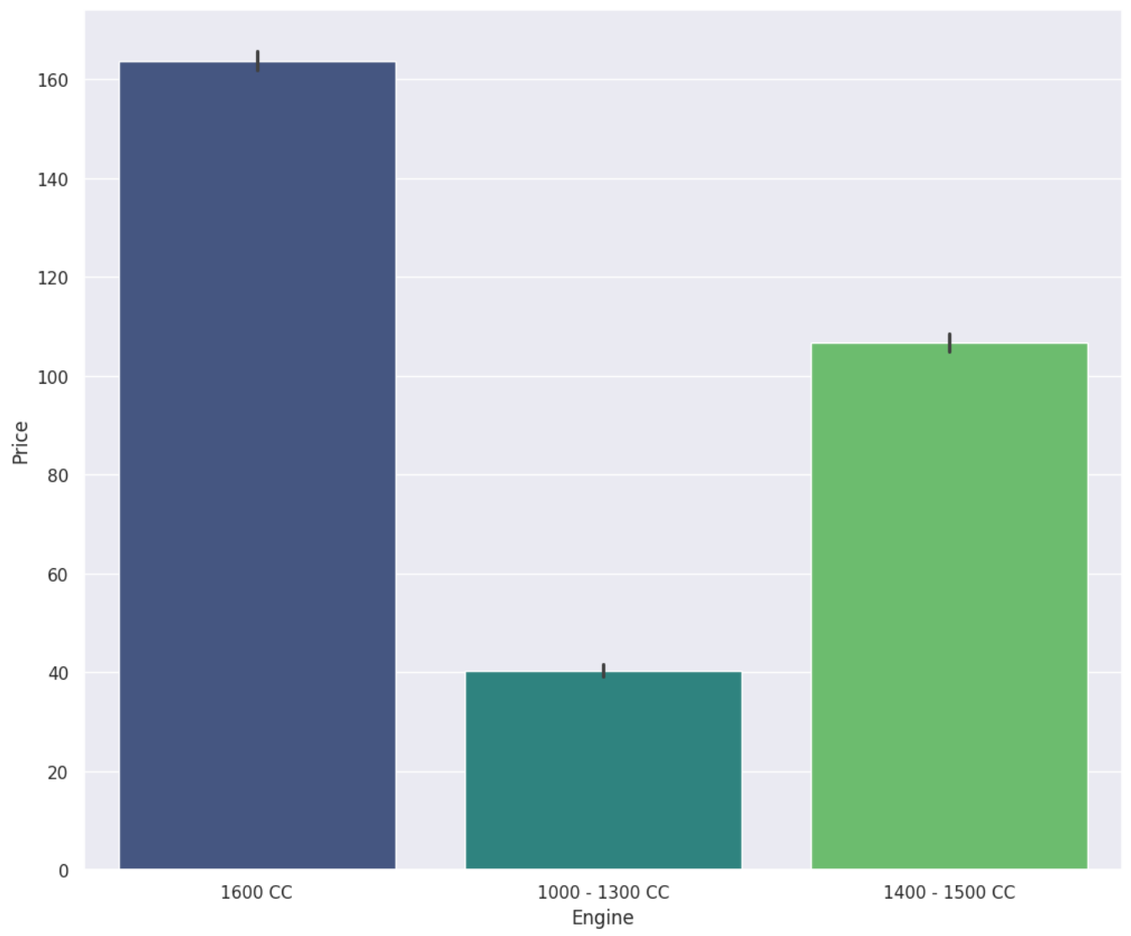


Distribution of cars by manufacturing year

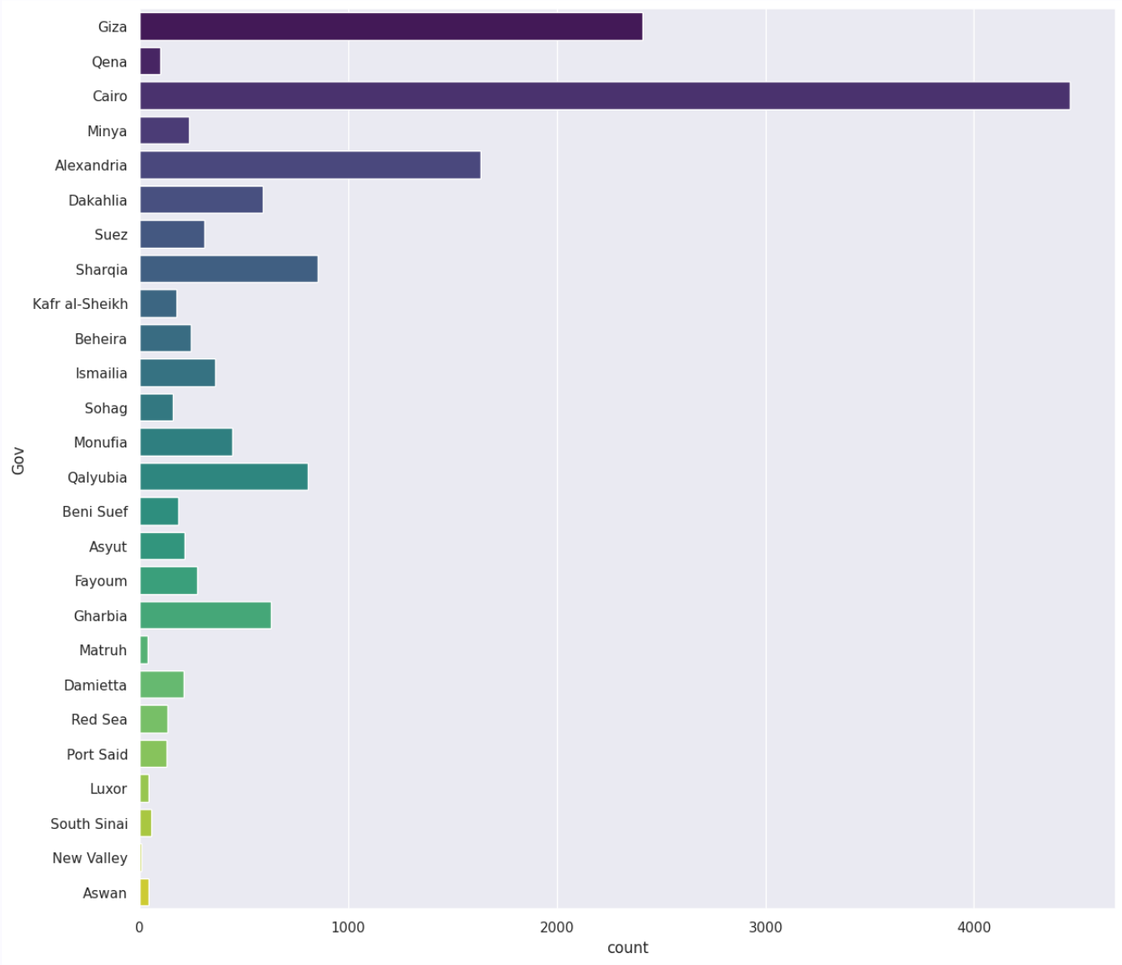


Exploratory Data Analysis (EDA)

Engine size distribution and impact on price

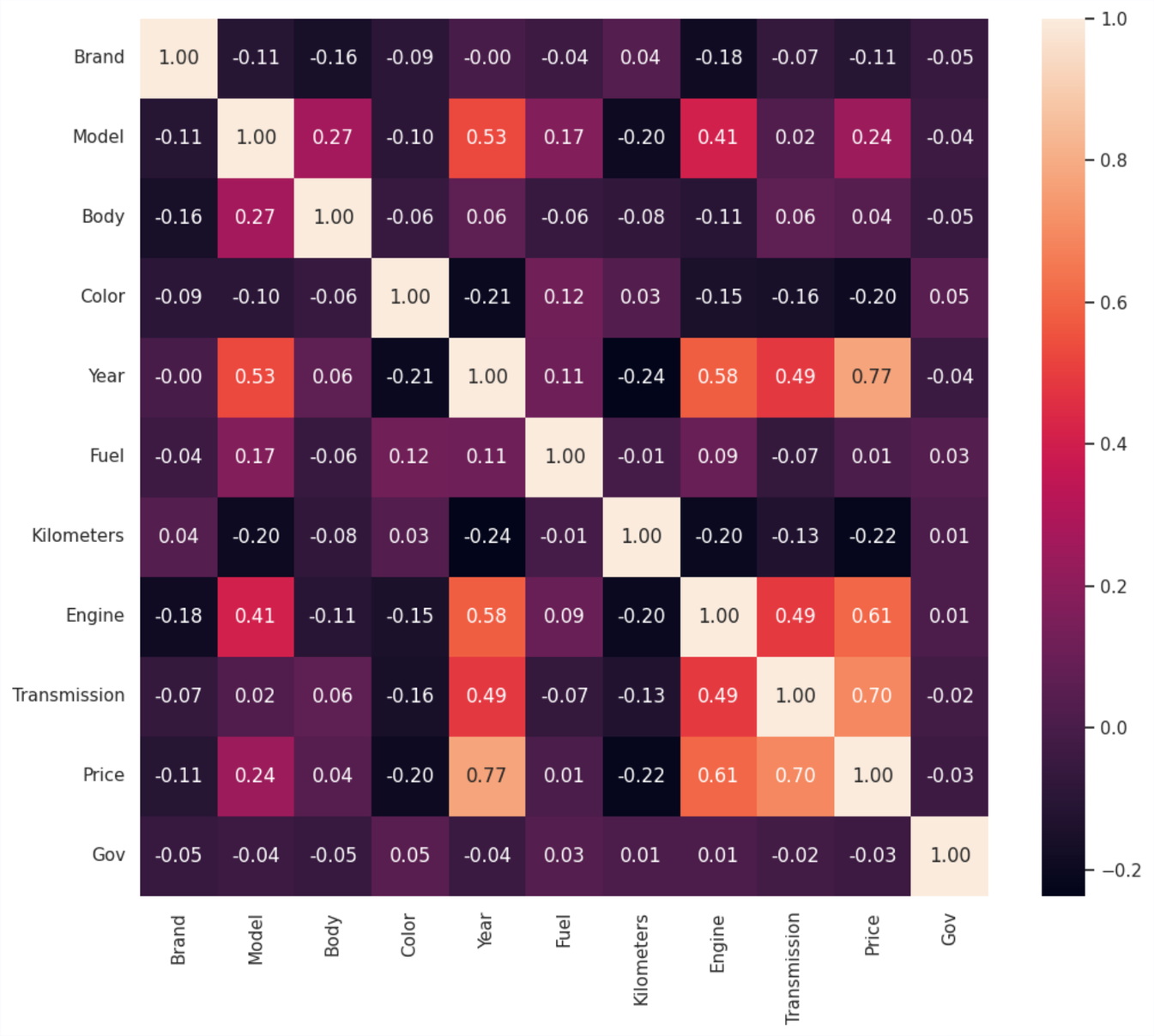


Distribution of cars by governorate (location)



Feature Selection

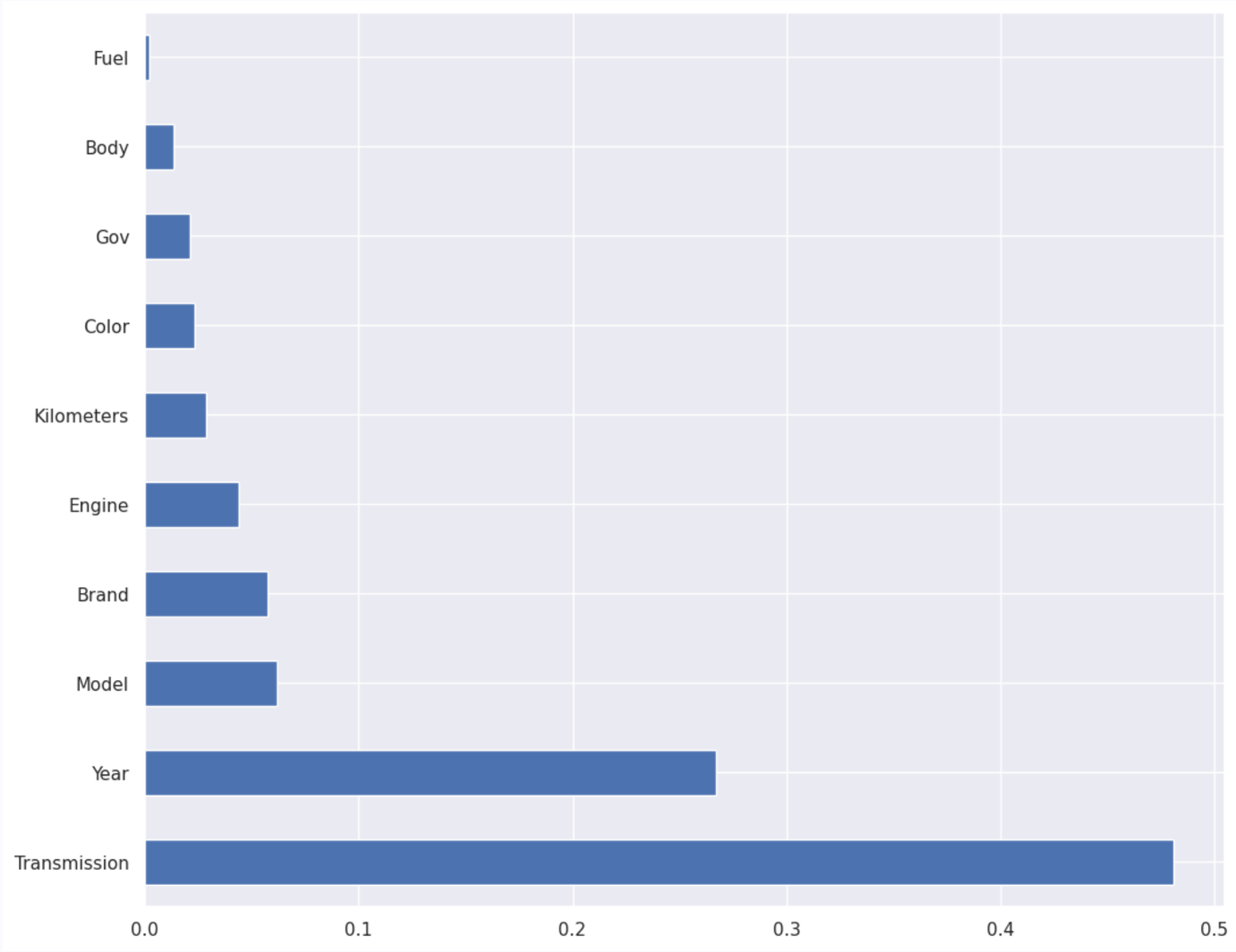
- Separate independent and dependent variables
- Display the correlation matrix between variables



Feature Selection

- Use Extra Trees Regressor to identify important features

- Split the data into features (x) and target (y).
- Visualize the correlation between features using a heatmap.
- Use Extra Trees Regressor to identify the most important features that influence the car price



Accuracy for Models

K-neighbors Regressor

The model achieved a high R2 score of **88.1%,** indicating its excellent ability to explain the variance in used car prices.

Decision Tree Regressor

The model achieved a high R2 score of **88.0%,** indicating its excellent ability to explain the variance in used car prices.

Random Forest Regressor

The model achieved a high R2 score of **91.6 %**indicating its excellent ability to explain the variance in used car prices.

XGB Regressor (Best Model)

The model achieved a high R2 score of **92.3 %,** indicating its excellent ability to explain the variance in used car prices.

Model Evaluation: R2 Score, MAE, and RMSE

R2 Score

The model achieved a high R2 score of **92.5%**, indicating its excellent ability to explain the variance in used car prices.

Mean Absolute Error (MAE)

The model's MAE of **13.36** suggests a low average error in its price predictions.

Root Mean Squared Error (RMSE)

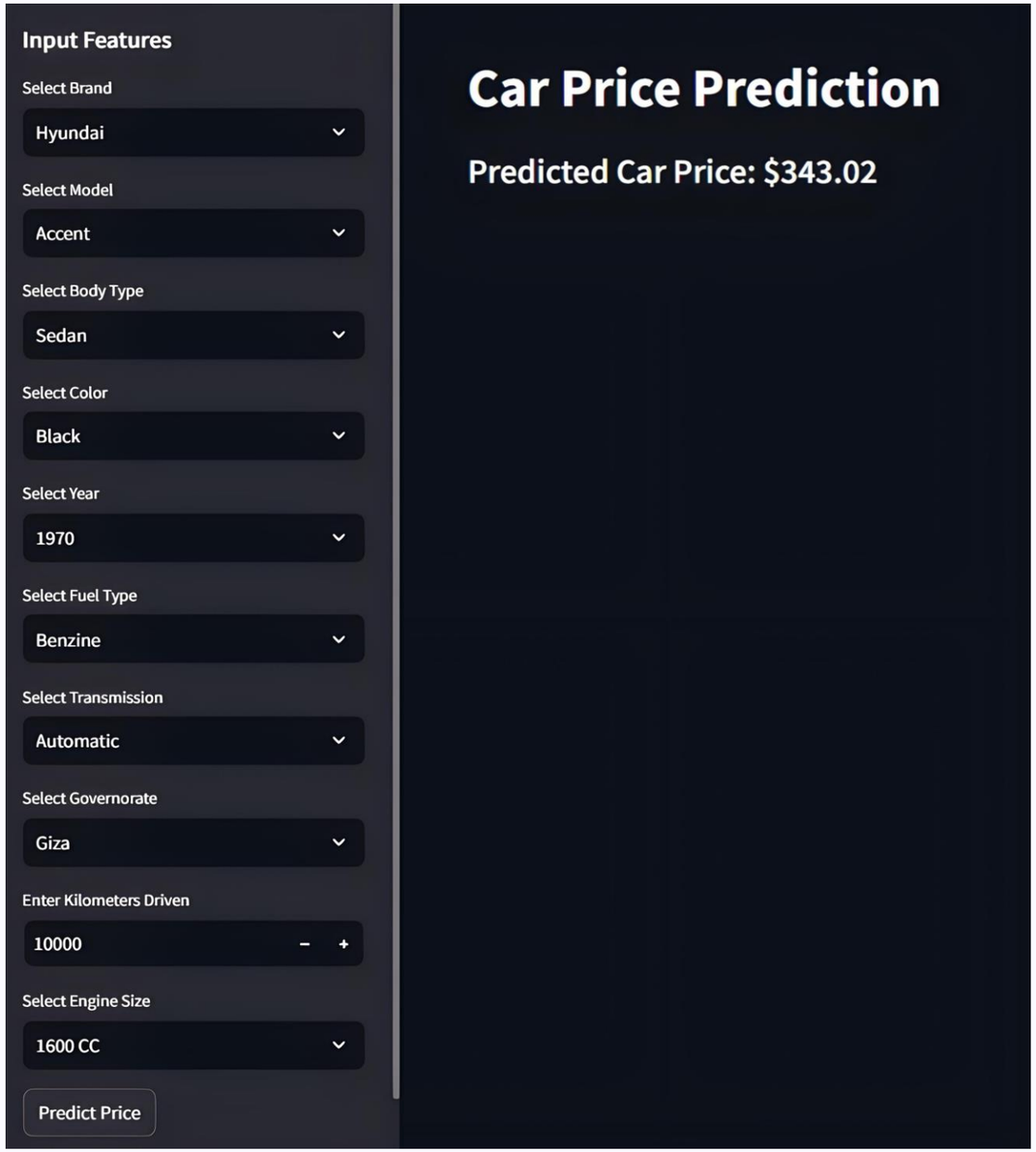
The RMSE **of 22.39** further confirms the model's high accuracy in forecasting used car prices.

Stream lit

This Stream lit app predicts car prices based on user-selected features. Simply choose options for:

- **Brand**
- **Model**
- **Body type**
- **Color**
- **Year**
- **Fuel type**
- **Transmission**
- **Governorate**
- **Kilometers driven**
- **Engine size**

[Notebook Link]



Thank You !