USED CAR PRICE PREDICTION

Agenda:

❖ Introduction

❖ Data Extraction

❖ Data Cleaning

❖ Exploratory Data Analysis

❖ Prediction Modelling

❖ Application Design

❖ Conclusion

Introduction

Objective:

The primary aim of this project is to predict the prices of used cars based on various features such as Type, Mileage, Performance & Features, etc.
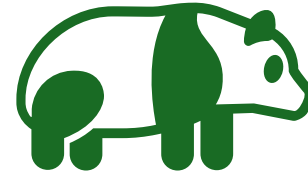
Motivation:

The prediction of used car prices is crucial for both buyers and sellers to make informed decisions. It helps sellers set competitive prices and assists buyers in assessing fair value.

Data Extraction

Source:

The dataset was sourced from CSV files containing unstructured data.

Processing:

Utilized Python libraries like pandas to transform raw CSV data into a structured format suitable for analysis.

Data Cleaning

- **Steps Taken:**

  - Handling missing values

  - Removing duplicates

  - Standardizing data formats

  - Addressing outliers

  - Encoding categorical variables

Exploratory Data Analysis

- **Key Insights:**
  - Statistical summaries of numerical features
  - Distribution plots
  - Correlation analysis between features and target variable (car prices)
  - Visualization of trends and patterns
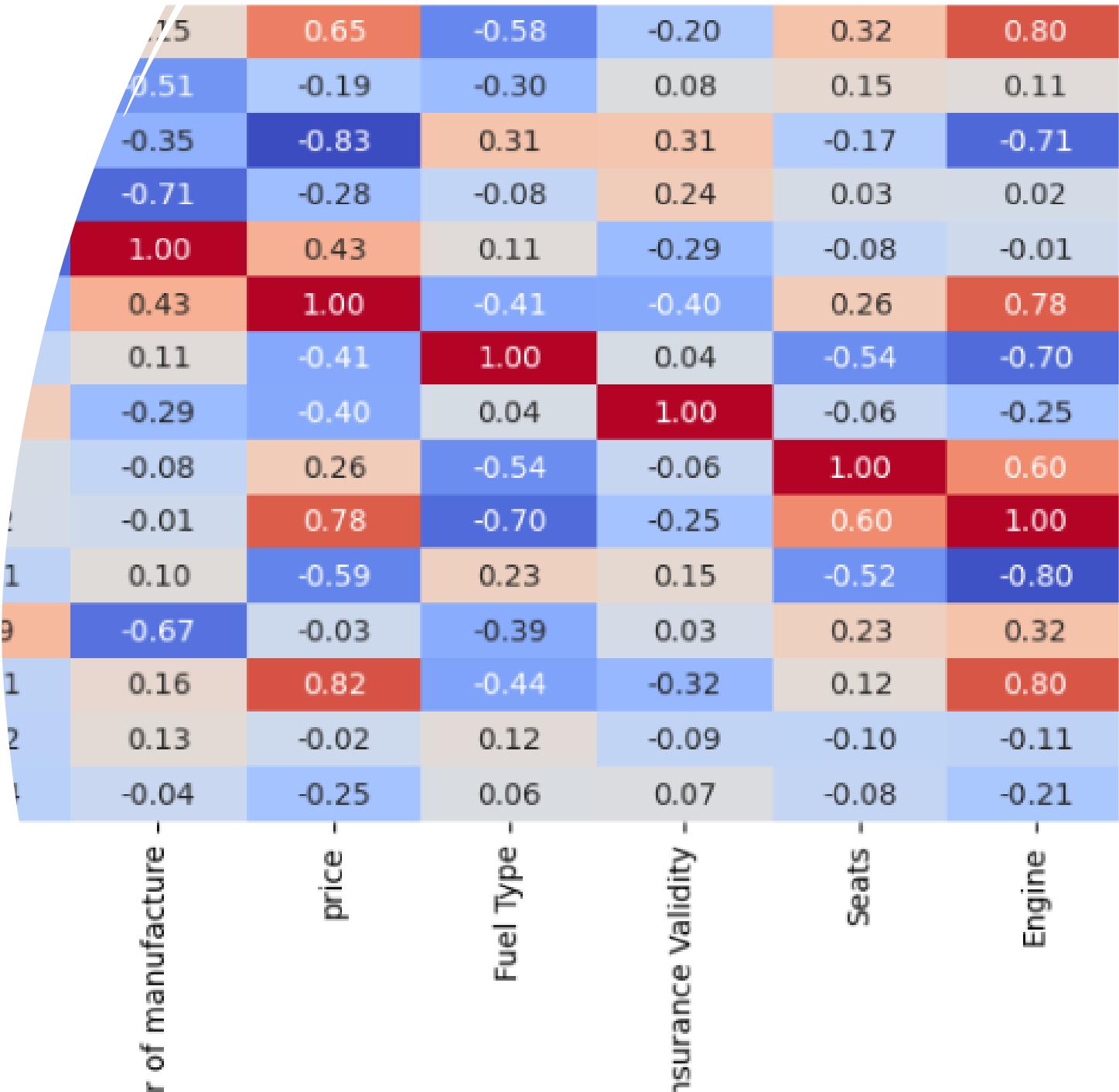
Prediction Modelling

Models Used:

Linear Regression

Random Forest

XG Boost

Gradient Boosting

Decision Tree Regressor

Evaluation:

Performance metrics such as MSE, MAE,R-squared, MAPE

Cross-validation to assess model robustness

Comparison of model accuracies

```
get_perform(train_lab, rfr_train_pred)
```

	mean_absolute_error	mean_squared_error	mean_absolute_percentage_error	r2_score
0	0.00137	0.000033	1.304036e+09	0.970811

```
get_perform(test_lab, rfr_test_pred)
```

	mean_absolute_error	mean_squared_error	mean_absolute_percentage_error	r2_score
0	0.003678	0.000124	0.166622	0.910688

- **Best Model Selection**

  Based on evaluation metrics, the Random Forest

  model emerged as the best performer due to its

  superior accuracy in predicting used car prices.

Application Design

- **Purpose:**
  - Developed a user-friendly web application using Stream-lit to allow users to predict car prices based on selected features.

- **Features:**
  - Input forms for user data
  - Integration with the deployed Random Forest model for real-time predictions
  - Visualizations to enhance user experience

<u>Conclusion</u>

- In this project, we successfully developed a model to predict used car prices using different machine learning algorithms.

- After comparing several models, Random Forest gave the best results.

- We also built a simple web app using Stream-lit, which allows users to easily predict car prices by entering key details.

- **Future Work**
  - Discuss potential enhancements or additional features that could be implemented to further improve the model's accuracy or usability of the application.

- **References**
  - Include information for datasets, libraries, and methodologies used throughout the project.
  - https://github.com/Syed-Abuthahir-M/Car-Dekho-Price-Prediction