  
**Project Title**

**Car Price Prediction**

Pushpendra Upadhyay (2215001372)

B.Tech. CSE 3rd year (Sec: 3E)

**Statement about the problem:** The used car market has seen a rise in demand, making it essential for sellers and buyers to determine the most appropriate price for a vehicle. Predictive analysis of car prices based on factors such as model, mileage, year, fuel type, and transmission is critical to provide accurate valuations.

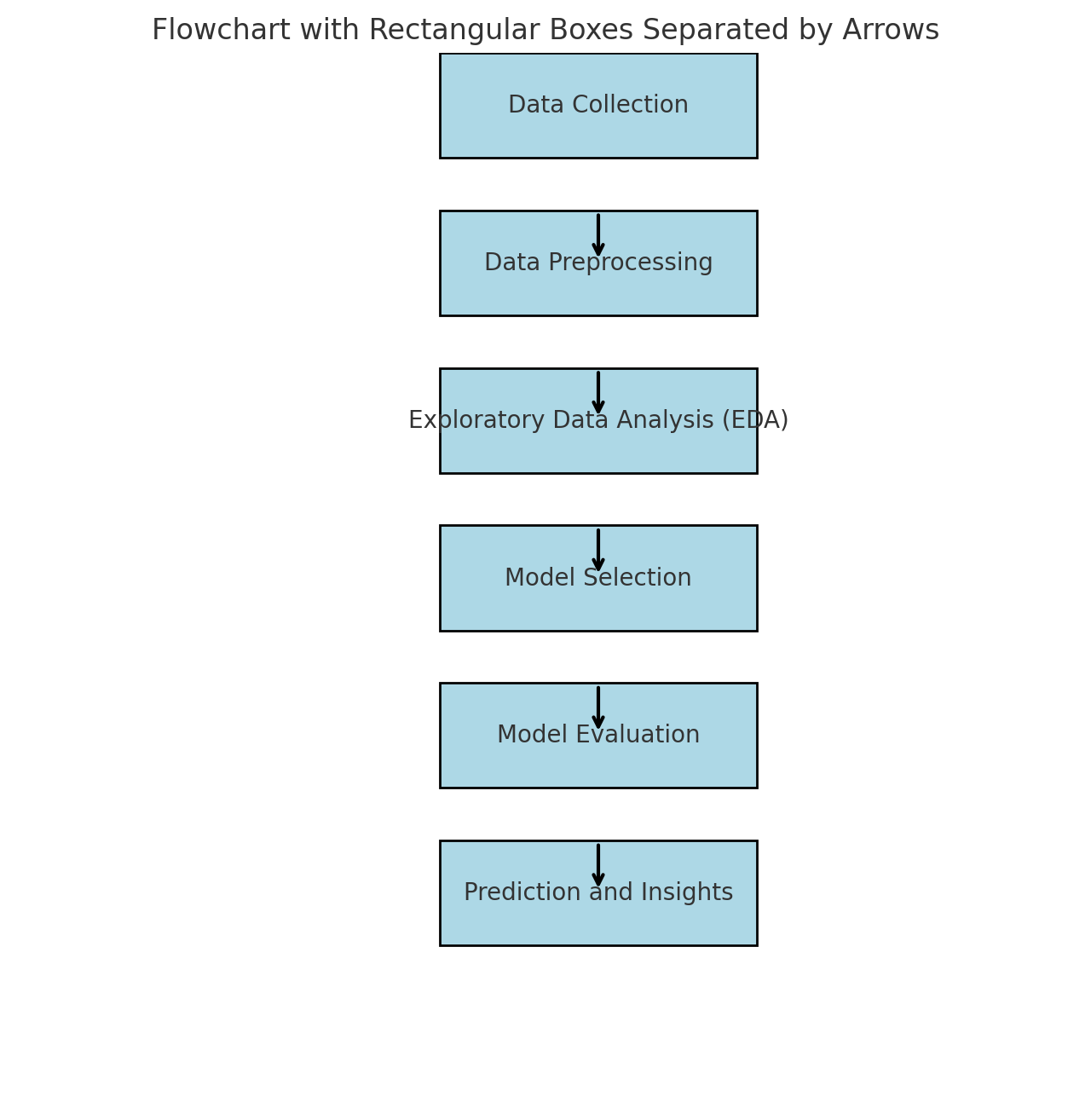
**Description:** The dataset contains cleaned information about used cars, including features like price, model, year, mileage, fuel type, and transmission. By performing data analysis and building a predictive model, the aim is to forecast the prices of used cars based on these features, providing a tool for both buyers and sellers in the car market.

**Objective and Scope of the Project:**

**Objective:** To analyse the given dataset and create a model that accurately predicts car prices based on various features like model, mileage, year, fuel type, and transmission.

**Scope:** The project involves data cleaning, feature selection, exploratory data analysis (EDA), and building a predictive model using machine learning algorithms such as Linear Regression. The model will be evaluated for performance, and predictions will be made on new data.

**Flow Diagram for this Model:**



1. **Data Collection:**

- Input: Cleaned car dataset.

- Output: Raw data ready for preprocessing.

1. **Data Preprocessing:**

- Handle missing values, outliers, and scale the data.

1. **Exploratory Data Analysis (EDA):**

- Visualization: Create plots and graphs to understand the distribution and relationships in the data

- Correlation Analysis: Identify correlations between features

1. **Model Selection:** Choose Algorithms: Select suitable machine learning algorithms (e.g., Linear Regression, Decision Trees, etc.)

- Model Training: Train the models using the training dataset.

1. **Model Evaluation:**

- Testing: Evaluate the model using the testing dataset

- Performance Metrics: Calculate accuracy, precision, recall, F1 score, etc.

- Model Tuning: Fine-tune hyperparameters for better performance

1. **Model Deployment:**

- Implementation: Deploy the model into a production environment (if applicable)

- Monitoring: Set up monitoring to track model performance over time

1. **Prediction and Insights:**

- Prediction: Use the model to make predictions on new data

- Insights: Generate insights from the model predictions for decision-making

**Methodology:** The process involves the following steps:

* Load the cleaned car data into the Python environment.
* Data Preprocessing: Clean the data, handle missing values, and prepare it for analysis.
* Exploratory Data Analysis: Visualize important relationships between car features and their prices.
* Model Training: Train models such as Linear Regression and Decision Trees using scikit-learn.
* Evaluation: Use metrics like R-squared, MSE to evaluate model performance.
* Visualizations: Use Python libraries like Matplotlib and Seaborn for graphs and charts.

**Hardware & Software used:**

* **Hardware:** Standard computing resources (laptop/desktop with adequate RAM and processing power).
* **Software:** Python, Jupyter Notebook, Pandas, Matplotlib, Seaborn, and other relevant Python libraries.

**Future Work of this Project:**

- Expanding the analysis to include global data.

-Integrating additional features such as car reviews, resale values, and brand effects.

- Developing a web-based tool or dashboard to predict car prices in real-time.

**References/Bibliography:**

* **scikit-learn Documentation:** [scikit-learn: machine learning in Python— scikit-learn 1.5.1 documentation](https://scikit-learn.org/stable/)
* **Python Official Documentation:** [Welcome to Python.org](https://www.python.org/)
* **Matplotlib library:** [Matplotlib — Visualization with Python](https://matplotlib.org/)
* **Seaborn library:** <https://seaborn.pydata.org>
* **Pandas library:** <https://pandas.pydata.org/>
* **Jupyter Notebook:** <https://jupyter.org/>