Cardheko Used Car Price Prediction

Overview

This project focuses on building a machine learning model to predict the prices of used cars based on various factors like car brand, model, year of manufacture, fuel type, transmission, and mileage. The aim is to help users understand market trends and make informed decisions when buying or selling used cars.

Project Structure

- **data/**: Contains the dataset used for training and testing.
- **notebooks/**: Jupyter notebooks for exploratory data analysis (EDA) and model development.
- **scripts/**: Python scripts for data preprocessing, model training, and evaluation.
- **models/**: Serialized trained models for use in predictions.
- **README.md**: Overview and instructions for the project.

Data Preprocessing

The dataset is cleaned and preprocessed by handling missing values, encoding categorical variables, and normalizing numeric features.

Model Training

Various machine learning algorithms are tested, including:

- Linear Regression
- Decision Trees
- Random Forest
- Gradient Boosting

Evaluation Metrics

- Mean Absolute Error (MAE)
- Mean Squared Error (MSE)
- Root Mean Squared Error (RMSE)
- R-squared Score

Results

The best-performing model is evaluated based on metrics and hyperparameter tuning.

Future Work

- Integrating deep learning models for improved accuracy.
- Creating a web application for real-time user predictions using frameworks like Flask or Streamlit.
- Expanding the dataset to include more cities and car brands.