Project Report

**Title**: Used Car Price Prediction Tool

**Project Overview**

**Objective**: The objective of this project is to develop an interactive and user-friendly web application using Streamlit to predict the prices of used cars based on several features, enhancing customer experience and streamlining the pricing process.

**Key Technologies**: Python, Streamlit, Scikit-learn, Pandas, NumPy

**Problem Statement**: Car Dheko wants to help customers and sales representatives estimate the price of a used car more accurately based on a wide range of features. The aim is to leverage machine learning to provide insights that streamline the pricing process and improve the customer experience.

**Target Audience**:

* **Customers**: To check if the price they are paying is reasonable.
* **Sales Representatives**: To assist in pricing negotiations with accurate data.

**Data Pre-processing**

* **Handling Missing Values**: the approach used for imputing missing values. Here we are removing more columns like similar columns and which affect data more
* **Encoding Categorical Variables**: we have used standard scaler for removing outliers and transforming it to fit into label encoder.
* **Feature Scaling**: Outline any normalization or scaling techniques applied to the features.
* **Handling Outliers**: If applicable, detail how you dealt with outliers in the data.
* **Final Features**: A list of features used in the final model, including those you selected to improve performance.

**Model Development**

* **Model Selection**: A brief description of the machine learning algorithms you tried:
  + Linear Regression
  + Decision Tree Regressor
  + Random Forest Regressor
  + Gradient Boosting Regressor
* **Model Evaluation**: Provide metrics like RMSE, MAE, R² for each model and justify why the Linear Regression was selected as the best model.
* **Hyperparameter Tuning**: here we are using grid search methods to fine-tune the model.

**Results**

* **Final Model Performance**: Summarize the performance of the final model (Linear Regressor) using metrics like RMSE, MAE, and R². finally, model was trained in linear regressor because it r2 score it accuracy and easy to train model
* **Benefits to Car Dheko**:
  + Improved customer satisfaction through more accurate price predictions.
  + Sales representatives can make data-driven decisions.