Used Car Price Prediction using Machine Learning

Team Members:

1. Prudvi Siddanathi – rmz769
2. Aanathi Mosali – bat761
3. Bhanu Anish Akkineni – umb160

Problem Statement

Accurately estimating the pricing of used automobiles is difficult owing to the numerous elements that determine a car's market worth, such as age, fuel type, gearbox, location, and ownership history. The lack of trustworthy and consistent pricing data causes considerable market differences, making it difficult for buyers to determine fair value and sellers to set competitive rates. This discrepancy reduces market transparency, confuses decision-making, and can lead to unjust pricing, with automobiles either undervalued or overpriced. There is a need for a strong, data-driven strategy to reliably predict used car prices that uses powerful machine learning algorithms to examine previous pricing data and determine the primary drivers of a vehicle's worth. This research intends to create and test prediction models using Machine Learning algorithms to provide accurate price estimations, thereby enhancing market transparency and fostering a more equitable automotive market.

Methods

- Linear Regression
- Random Forest
- SVM
- Decision Tree
- XGBoost

Dataset and Evaluation Metrics

Dataset - Historical data on used cars from Cars 24 (acquired from Kaggle).

Evaluation Metrics:

- Mean Squared Error (MSE)
- Root Mean Squared Error (RMSE)
- R-squared (R^2) Score