Predicting the Prices of Used Cars with Machine Learning

Aarya Gupta
IIIT Delhi
aarya22006@iiitd.ac.in

Aditya Raj Jain
IIIT Delhi
aditya22037@iiitd.ac.in

Adarsh Jha
IIIT Delhi
adarsh22024@iiitd.ac.in

Krishna Shukla
IIIT Delhi
krishna22254@iiitd.ac.in

I. Motivation

The factors affecting prices in the used car market are complex, including make, model, year of manufacture, mileage, and condition. Accurately determining the price has been challenging for both sellers and buyers. Traditional methods often fail to capture these intricacies, leading to inefficiencies and inaccurate valuations. This project aims to develop a machine learning model to predict used car prices more accurately, providing a reliable tool to help users make informed decisions in this dynamic market.

II. Background of the Problem

The second-hand car market has expanded significantly online, complicating vehicle pricing. Factors like age, mileage, brand, and condition affect value, but traditional methods often use basic heuristics or subjective judgment, causing inconsistent pricing.

This project proposes a data-driven solution using machine learning to improve price predictions. By analyzing large datasets, machine learning uncovers hidden patterns, ensuring more accurate, fair pricing, benefiting both buyers and market efficiency.

III. Scope

The project will develop a platform with a user-friendly interface that predicts prices for used cars based on essential attributes such as make, model, year, mileage, and vehicle condition.

IV. Features

- **Data Collection & Processing:** Collect and process data to ensure accuracy and completeness.
- **Model Comparison:** Build and test various models (e.g., Linear Regression, Random Forest, XGBoost) to select the best-performing model.
- **Real-time Updates:** Provide real-time predictions and trends based on the latest data.
- **Additional Features:** Offer personalized recommendations for car purchases and regular updates on model performance and accuracy.

V. Timeline

- **Week 1-2:** Define project objectives and collect initial data.
- **Week 3-4:** Clean and preprocess the data; perform feature engineering.
- **Week 5-6:** Build and test different machine learning models.
- **Week 7:** Compare model performance and select the best model.
- **Week 8-9:** Integrate the selected model into the platform and develop additional features.
- **Week 10:** Final testing, evaluation, and deployment of the model.

VI. Individual Tasks

Aarya Gupta: Data Collection and Preprocessing

- Collect data from various sources and preprocess it to ensure consistency and accuracy.
- Handle missing values, outliers, and perform initial data analysis.

Aditya Raj Jain: Model Development and Training

- Detect highly correlated features.
- Research & implement various machine learning models.
- Train models using historical data and fine-tune parameters for optimal performance.

Adarsh Jha: Model Evaluation and Feature Engineering

- Develop and select relevant features for model training.
- Evaluate model performance using appropriate metrics and refine the model for optimal results.

Krishna Shukla: Real-time Updates and personalized recommendations

- Develop real-time predictions created using the information available on the platform, and trends received worldwide.
- Personalized recommendations for car purchasing and regular updates on model performance.

VII. Final Outcome

Expected Results:

- A robust machine learning model can accurately predict used car prices.
- A platform that provides real-time predictions and personalized recommendations.
- Enhanced user experience in the used car market through data-driven decision-making.

Contribution:

- Provide a reliable, data-driven tool for pricing used cars in the Indian market.
- Improve transparency and fairness in the used car market by leveraging machine learning techniques.

VIII. References

- Used Car Price Prediction Dataset : It contains most all relevant information that Craigslist provides on car sales including columns like price, condition, manufacturer, latitude/longitude, and 18 other categories
- Research Paper 1
- Research Paper 2
- Research Paper 3