Used Car Price Prediction - Final Report

Audience

Used car dealers seeking to optimize pricing strategies and understand what factors most influence vehicle value.

Objective

To build a data-driven model that predicts used car prices based on features such as mileage, age, manufacturer, fuel type, transmission, and body type. The goal is to uncover the most influential factors and support better pricing and inventory decisions.

Key Findings

- Car Age - Newer cars retain significantly more value.
- Mileage (Odometer) - Higher mileage strongly reduces price.
- Manufacturer - Brand reputation matters (e.g., Toyota, Ford, BMW).
- Vehicle Type - SUVs and trucks are priced higher than sedans.
- Fuel Type & Transmission - Gasoline and automatic vehicles dominate pricing patterns.

Model Summary

Model Used: Gradient Boosting Regressor (Tuned)

R2 Score: 0.7687 (model explains ~77% of price variability)

RMSE: $6,722 (average prediction error)

MAE: $4,281 (half of predictions within this margin)

This model is accurate enough for practical use in pricing or appraisal tools.

Visual Insights

- Actual vs. Predicted - Strong alignment indicates reliable pricing predictions.
- Residuals - Errors are normally distributed; no major systemic bias.
- Feature Importance - Confirms key business insights: age, mileage, brand matter most.

Recommendations for Dealers

- Prioritize newer, low-mileage vehicles when acquiring inventory.

- Use manufacturer and type insights to target high-margin segments.
- Implement this model into your pricing tools or websites for real-time valuations.
- Use insights for marketing (e.g., 'lower mileage = better deal!').

Next Steps (Optional Enhancements)

- Build a live dashboard or API for pricing.
- Integrate more market-specific features (e.g., region, seasonality).
- Use SHAP or LIME for explainable AI to show why a car is priced a certain way.