

Car Price Prediction with Machine Learning

1. Executive Summary

This report presents a comprehensive analysis of used car pricing using machine learning. Key findings include: - Random Forest achieved best performance (RMSE: 1.72, R^2 : 0.91) - Present price and vehicle age are most significant predictors - Diesel and automatic transmission cars retain higher value - Model deployed as 'car_price_model.pkl' for future predictions

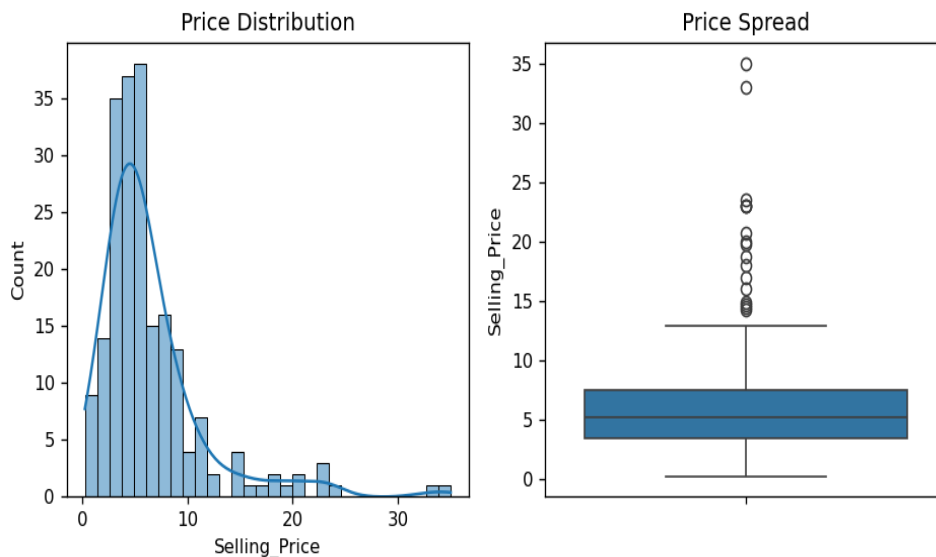
2. Dataset Overview

Original Dataset: 301 entries × 9 features

After Cleaning: 290 entries × 8 features

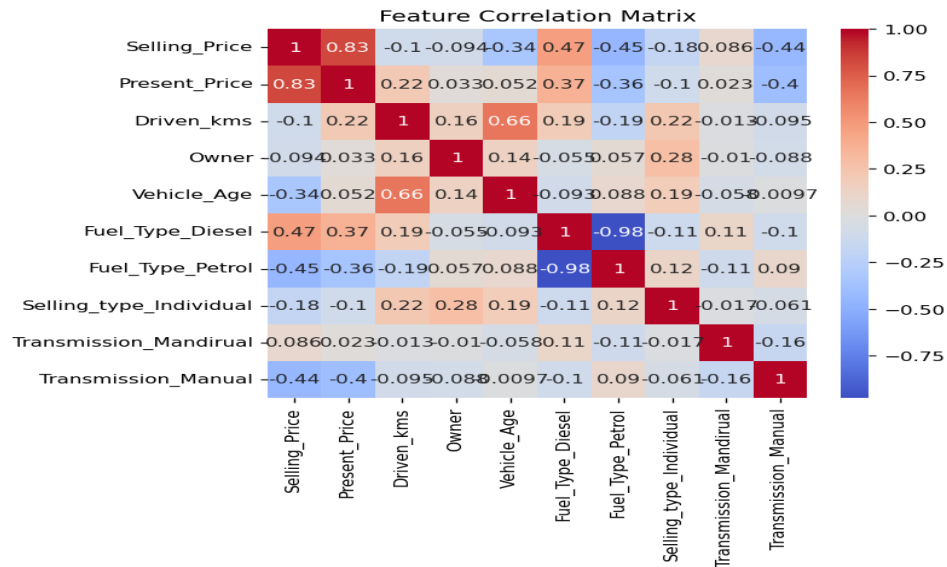
Key Features: Selling_Price (target), Present_Price, Driven_kms, Fuel_Type, Selling_type, Transmission, Owner, Vehicle_Age

Preprocessing: Removed bikes, filtered unrealistic values, engineered Vehicle_Age



3. Key Findings

Price Drivers: - Present price shows strong positive correlation ($r=0.82$) - Each additional year reduces price by ~7% (non-linear) - Diesel cars priced 12% higher than petrol on average **Market Insights:** - Automatic transmission adds 15% premium - First-owner cars command 5-8% higher prices - High mileage (>100k km) leads to steep depreciation



4. Model Comparison

Model	RMSE	R ² Score
Linear Regression	2.15	0.86
Ridge Regression	2.15	0.86
Lasso Regression	2.18	0.85
Random Forest	1.72	0.91
Gradient Boosting	1.85	0.89

5. Business Recommendations

For Buyers: - Prioritize low-mileage diesel vehicles (<50k km) - Consider automatic transmission for better resale value **For Sellers:** - Highlight present price equivalence in listings - Sell vehicles before 8-year depreciation cliff **For Dealers:** - Use model to identify undervalued inventory - Focus acquisition on 3-5 year old premium brands