

PROJECT ABSTRACT

Used Car Price Prediction using Gradient Boosting

A gradient-boosting regressor estimating fair used-car prices from listing attributes, with an explainable price estimator and market-trend analysis.

Introduction / Background

India's used-car market runs on information asymmetry: sellers anchor high, buyers guess low, and the fair price of a specific car with its mileage and history is folklore. Listing platforms show asking prices, not transaction prices. A model trained on thousands of real listings replaces folklore with an estimate: gradient-boosted trees learn how age, mileage, fuel, transmission, brand and ownership move prices, returning a fair value with an expected band.

Explainability is what makes it usable - the demo breaks down each factor's contribution, so a buyer can negotiate from evidence rather than instinct.

Problem Statement

Buyers overpay or walk away from fair deals; sellers misprice and wait months. Existing price guides use coarse depreciation tables that ignore the specific car's attributes.

Objectives

Train a gradient-boosting regressor on ~9,200 listings with categorical handling. Evaluate with R² and MAPE plus residual analysis. Build an estimator demo giving price, band and per-factor breakdown, with listing verdicts and depreciation curves.

Proposed Methodology

Listings are deduplicated, outliers trimmed, and features engineered: vehicle age, log-mileage, brand, segment, fuel, transmission, owner count, city tier. A LightGBM-style regressor trains with cross-validation; held-out listings give R², MAPE and residual plots. Partial dependence reveals the depreciation structure.

System Architecture / Working

Entered specs flow through the fitted depreciation structure to produce a price in lakh rupees, a $\pm 10\%$ expected band, and the per-factor contribution bars. Sample listings are scored to fair/overpriced/suspicious verdicts with the buyer's playbook.

Hardware / Software Requirements

Hardware: laptop; training takes minutes on CPU. Software: Python 3, LightGBM/XGBoost, scikit-learn, pandas, NumPy, Matplotlib; the demo is plain HTML/CSS/JavaScript.

Expected Outcomes

A regressor reaching the R² 0.93 / MAPE 8% design targets with an estimator demo buyers could genuinely use and a negotiation guide in the report.

Applications

Car-buying assistants, dealership pricing tools, classified platforms adding fair-price badges, and as a template for any price-estimation task.

Future Scope

Image-based condition assessment, transaction-price (vs asking-price) calibration, and live listing ingestion with verdict APIs.