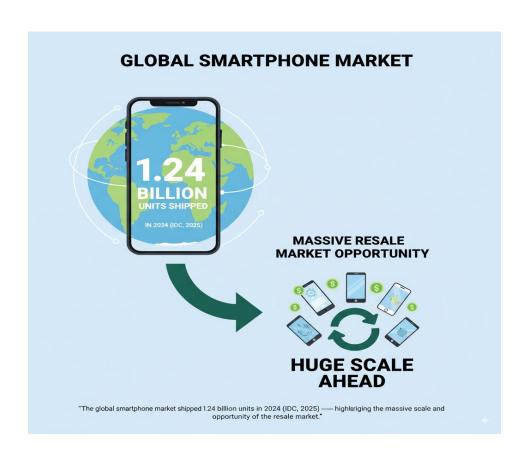


 Fair Prices · Smarter Decisions · Stronger Trust.

Professor: Dr. Willie Rivers

BY PHANIDHAR KASUBA

INTRODUCTION



Objective:

- Analyze used smartphone data to understand what drives resale value.
- Build predictive models to set fair, competitive resale prices.

Dataset Overview:

- 3,454 records from multiple smartphone brands & models.
- 15 attributes including:
 - Device Specs: RAM, Storage, Screen Size, Battery, Cameras, Weight
 - Usage Factors: Days Used, Release Year
 - Connectivity: 4G, 5G, OS, Brand
 - Prices: Normalized New Price, Normalized Used Price

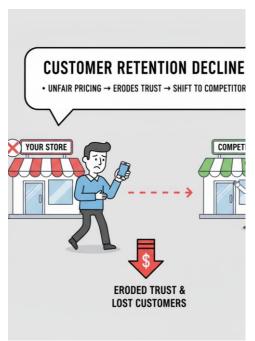






Undervaluation = Lost Revenue

BUSINESS PROBLEMS











BUSINESS & ANALYTICAL GOALS

BUSINESS GOALS



Optimize resale pricing to reduce inventory backlog



Maximize margin on high-value devices



Improve trade-in offers to increase turnover volume



Build customer trust & retention with transparent pricing



PROFIT & LOYALTY

A transparent, data-driven method for fair pricing drives sales, safgurads margins, and builds customer confidence.

ANALYTICAL GOALS



Low Error & High Accuracy

Price predictor & tiering clasfifier.



Stable & Robust

Across brands, OS, time and data issues



Clear Explanations

Understand pricing reaning.



Practical Diagnostics

Levers: memory, camera, age, network, etc



Well Calibrated

Predicted probabilities match observed levels



Measure Uncertainty

Prediction intervals for for guardrails

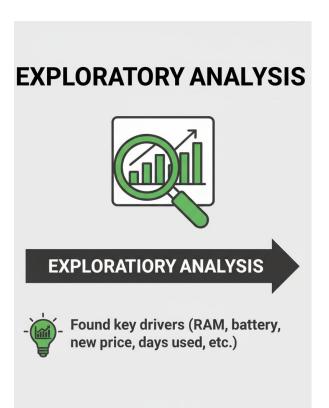


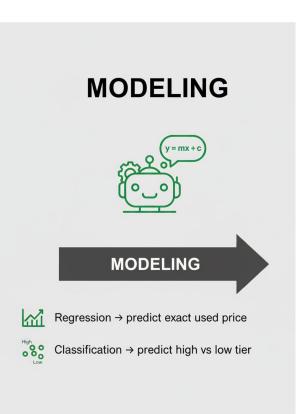
ANALYTICAL APPROACH

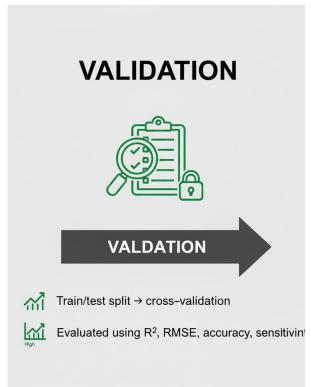


DATA PREPARATION

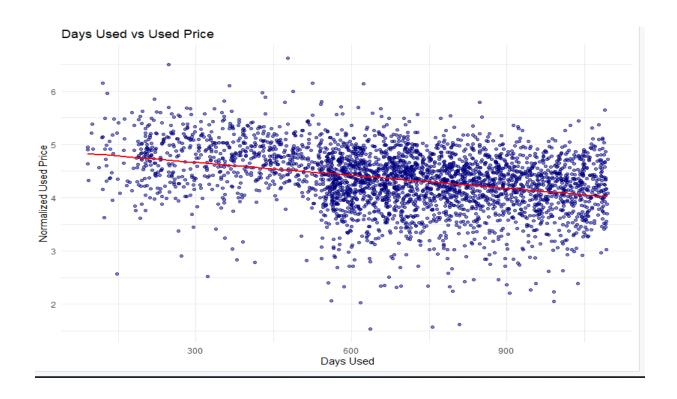
- Cleaned missing values, handled zeros, treated outliers
- Defined predictors (brand, OS, specs, usage usage)

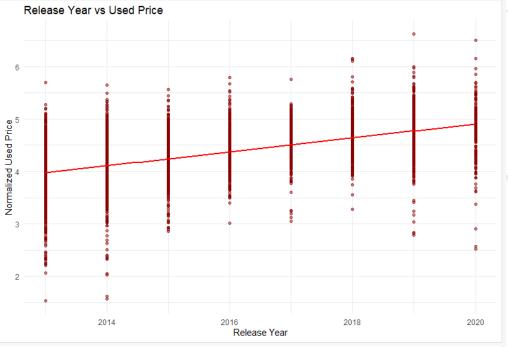


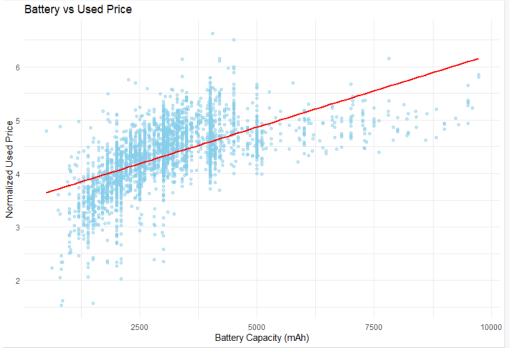


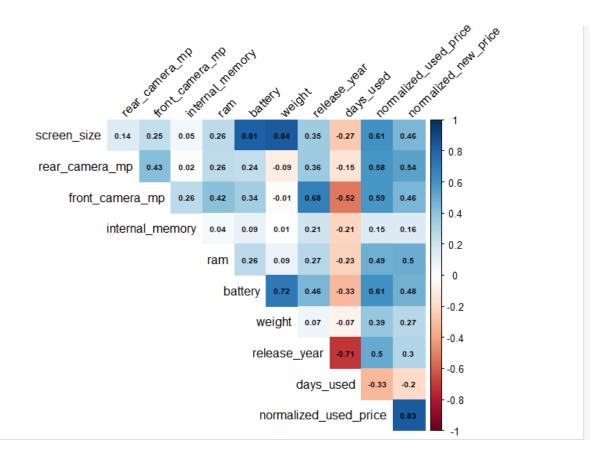


KEY VISUALIZATIONS

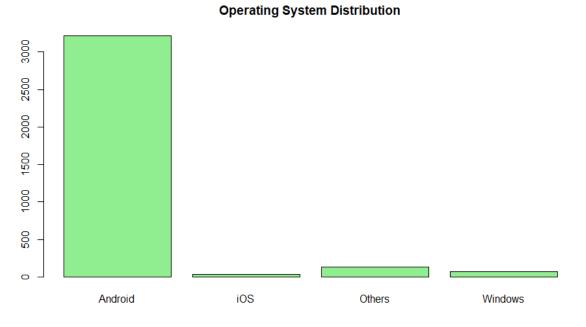












REGRESSION MODEL – PREDICT THE USED DEVICE PRICE

- Random Forest Regression delivered the best fit ($R^2 = 0.860$, RMSE = 0.226) \rightarrow most reliable for predicting fair resale prices.
- ➤ New Price is the strongest driver → higher original cost directly translates to higher resale value, reducing undervaluation risk.
- ➤ RAM, Battery, and Release Year are key value adders → newer devices with higher RAM and stronger batteries retain more market value.
- ➤ Days Used strongly impacts depreciation → longer usage sharply lowers resale price, guiding trade-in offers.
- ➤ Interpretability vs Accuracy trade-off → Linear/regularized models explain drivers clearly, while ensemble models like Random forest provide maximum pricing accuracy.

| Model | R ² (High RMSE(Low= =Better) Better) | |
|------------------------|--|-------|
| Linear Regression | 0.849 | 0.235 |
| Stepwise Regression | 0.846 | 0.238 |
| Random Forest (Reg) | 0.860 | 0.226 |

CLASSIFICATION MODEL – CLASSIFY THE DEVICES INTO HIGH OR LOW TIER

- Logistic Regression achieved 87.9% accuracy with balanced sensitivity (0.896) and specificity (0.858) – Highly interpretable, showing how predictors like RAM, release year, and connectivity features influence device value. Best when transparency is needed for stakeholders.
- ➤ Decision Tree reached 85.6% accuracy with clear rule-based splits (e.g., new price, battery, camera resolution) Intuitive to explain, useful for business adoption, though slightly less accurate and prone to overfitting on complex data.
- ➤ Random Forest delivered 87.6% accuracy with the highest sensitivity (0.913) Most robust and dependable, capturing complex feature interactions and minimizing variation. Best when predictive power and reliability are prioritized, though less interpretable than simpler models.

| Model | Accura cy | F1 Score | Sensitiv ity TP rate | Specifici ty TN rate |
|------------------------|--------------|-------------|----------------------------|----------------------------|
| Logistic Regression | 0.879 | 0.890 | 0.8961 | 0.8576 |
| Decision Tree | 0.856 | 0.871 | 0.8933 | 0.8102 |
| Random Forest | 0.876 | 0.889 | 0.9129 | 0.8305 |

CONCLUSION & RECOMENDATIONS

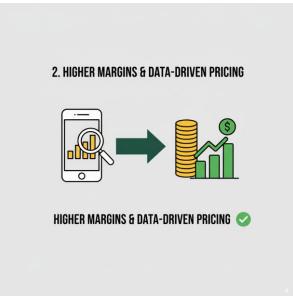
- > Pricing is multi-factor & dynamic manual methods can't keep pace with today's market.
- ➤ Key value drivers confirmed Screen size, RAM, battery, release year, connectivity; days used drives depreciation.
- > Best price predictor: Random Forest (Reg.) & accurate classifiers are Logistic Reg. & Random Forest
- > Our hybrid approach balances accuracy with transparency, giving you a scalable, defensible pricing system that improves margins and customer retention.

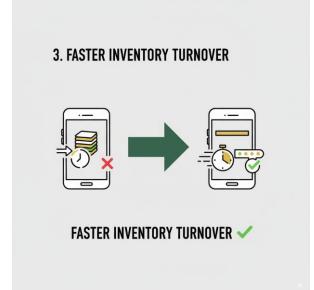
Recommendation

Deploy Random Forest for production pricing & auto-tiering and **use Logistic/Linear/Tree** to **explain** prices to stakeholders and customers

BUSINESS IMPACT

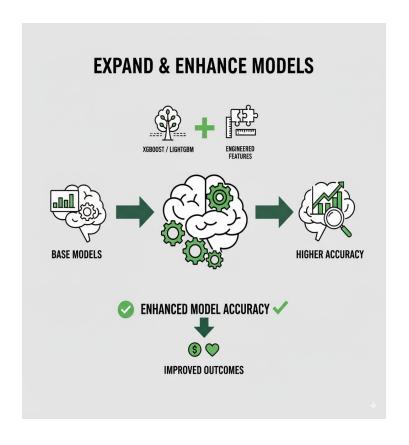




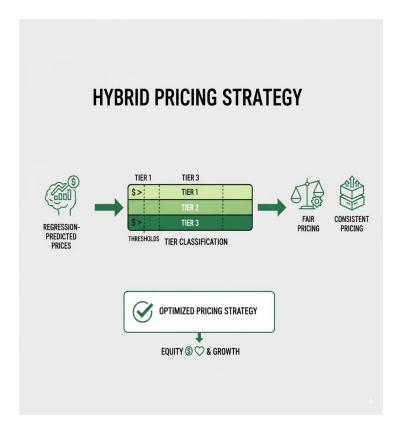




FUTURE WORK







Thank you