UC San Diego

Used Phones & Tablets Prices Prediction

Team 3 - Eric Gu, Hongjie Wang, Hesam Mojtahedi , Ching-Hao Wang, Chainika Shah

ECE 143: Programming for Data Analysis | March 12, 2025

Contents

1 Introduction - Problem Statement

2 Data Exploration & Visualization Web Demo

3 Price Prediction Using Multiple Models

4 Results and Key Observations

5 Conclusion and Further Scope

1 Introduction - Problem Statement

Understanding the Problem

- **Growing Market**: The used and refurbished device market has expanded significantly due to cost-effectiveness for consumers and businesses.

- **Pricing Challenge**:
  - Too low → Financial losses for sellers
  - Too high → Poor sales performance
  - Consumers struggle to determine fair prices

- **Need for a Solution:**
  - A data-driven approach is essential to accurately predict used phone and tablet prices.

Contents

2 Data Exploration & Visualization Web Demo

Dataset

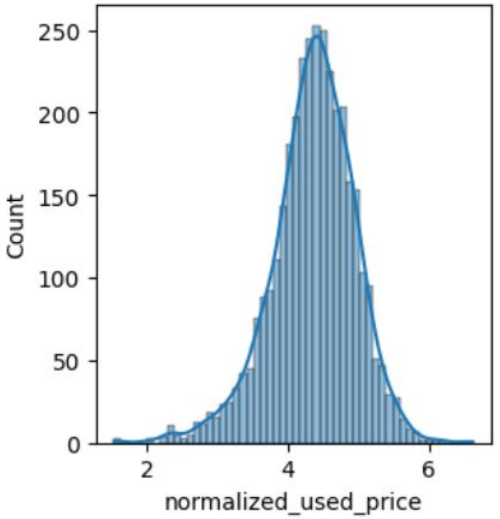
https://www.kaggle.com/datasets/ahsan81/used-handheld-device-

UC San Diego

3454 entries with 15 features each

used_device_data

device_brand	os	screen_size	4g	5g	rear_camera_mp	front_camera_mp	internal_memory	ram	battery	weight	release_year	days_used	r
Honor	Android	14.5	yes	no	13	5	64	3	3020	146	2020	127	
Honor	Android	17.3	yes	yes	13	16	128	8	4300	213	2020	325	
Honor	Android	16.69	yes	yes	13	8	128	8	4200	213	2020	162	
Honor	Android	25.5	yes	yes	13	8	64	6	7250	480	2020	345	
Honor	Android	15.32	yes	no	13	8	64	3	5000	185	2020	293	
Honor	Android	16.23	yes	no	13	8	64	4	4000	176	2020	223	
Honor	Android	13.84	yes	no	8	5	32	2	3020	144	2020	234	
Honor	Android	15.77	yes	no	13	8	64	4	3400	164	2020	219	
Honor	Android	15.32	yes	no	13	16	128	6	4000	165	2020	161	
Honor	Android	16.23	yes	no	13	8	128	6	4000	176	2020	327	
Honor	Android	15.47	yes	no	13	8	64	3	3020	150	2020	268	
Honor	Android	15.32	yes	no	13	8	64	4	5000	185	2020	344	
Honor	Android	16.69	yes	yes	13	16	128	8	4100	206	2019	537	
Honor	Android	15.32	yes	no	13	16	64	4	4000	171.5	2019	336	
Honor	Android	14.5	yes	no	13	5	32	2	3020	146	2019	230	



 # Mean: 4.36 corresponding to USD$318

Data Analysis and Visualization

UC San Diego

- Dataset Web Demo: https://wi25-ece143-team3.streamlit.app/

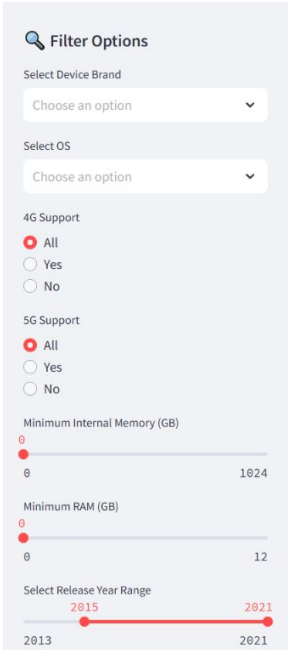
Main Features:

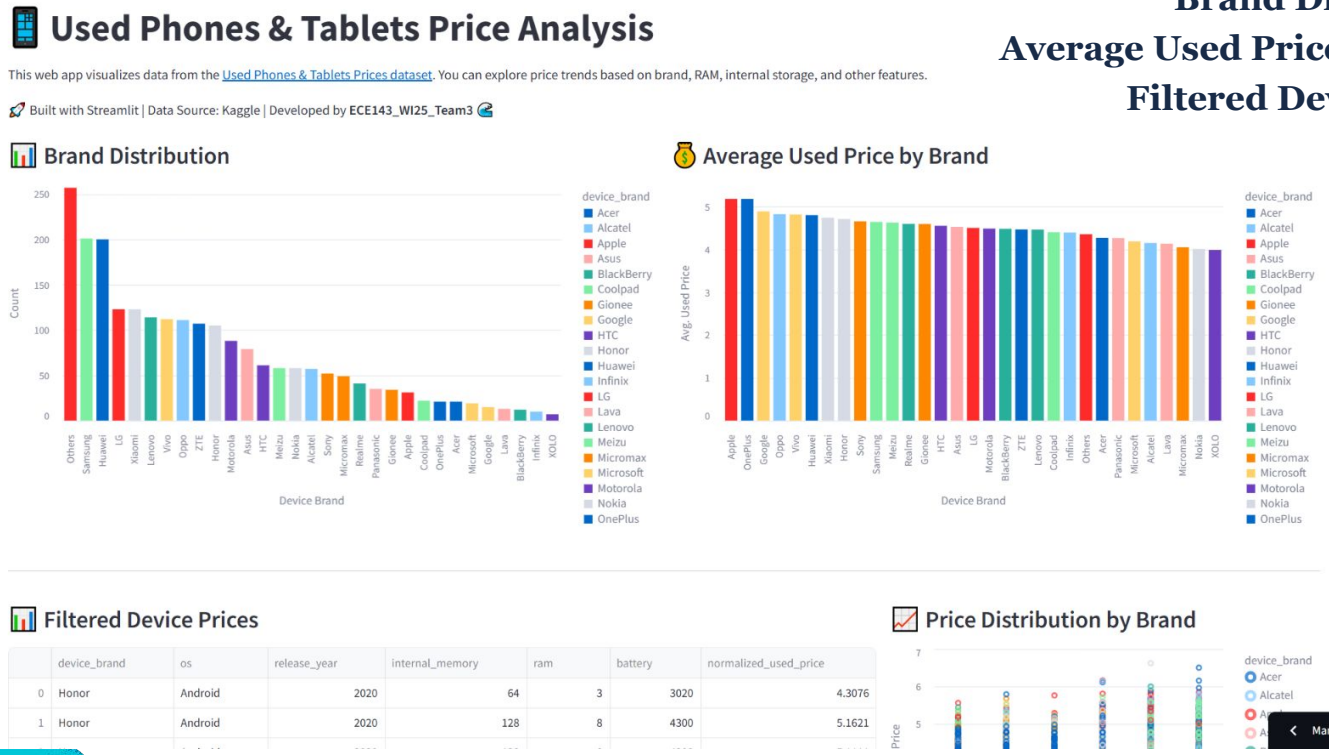
Brand Distribution

Average Used Price by Brand

Filtered Device Prices

...





	device_brand	os	release_year	internal_memory	ram	battery	normalized_used_price
0	Honor	Android	2020	64	3	3020	4.3076
1	Honor	Android	2020	128	8	4300	5.1621

3 Price Prediction Using Multiple Models

Price Prediction - Regression Problem

- We explored multiple models to predict used device prices:
  - **Linear Regression** – Established a baseline performance
  - **Neural Networks** – Captured complex patterns in data
  - **K-Nearest Neighbors** (**KNN**) – Based on similarity to past data points
  - **Random Forest** – Bagging learning for improved accuracy
  - **Support Vector Machine** (**SVM**) – Effective for high-dimensional data
  - **XGBoost** – Optimized gradient boosting for superior performance

4 Results and Key Observations

Results

Key Findings:

- **The original price of a device is the strongest predictor**
  - This aligns with market trends: high-end phones hold their value better.
- **Tree-based models (Random Forest, XGBoost) and SVM performed best**, with SVM achieving the **lowest RMSE of 0.3011**.
- **KNN showed surprising efficiency**, indicating that price prediction can benefit from finding similar past sales.
- **Neural network could improve with more data points**, but struggled slightly compared to tree-based models.

Model	RMSE	MAE
Linear Regression	0.371775	0.289166
NN Model	0.320985	0.247323
KNN Model	0.307225	0.233728
Random Forest	0.301147	0.233340
SVM Model	0.301137	0.228863
XGBoost Model	0.306823	0.239365

5 Conclusion and Further Scope

Conclusion and Further Scope

Key Takeaways and Future Work

- **Data Freshness is Critical** – The smartphone market evolves rapidly, so models need frequent updates.
- **Enhancing our Web Demo** – Future work could allow users (sellers) to enter phone details and get an estimated price.
- **Feature Engineering** – Exploring additional data like phone condition, repair history, and market trends could further improve predictions.
- **Hyperparameter Tuning** – Optimizing models like XGBoost for even better accuracy.

UC San Diego

Thank you!