

LAPTOP PRICE PREDICTION USING REGRESSION TECHNIQUES

(LEVERAGING DATA SCIENCE FOR ACCURATE PRICE ESTIMATION)



HAIDER ALI ADEEL – 2024MSDS-116

PROBLEM DEFINITION



Problem Statement:

The goal of this project is to predict laptop prices based on technical specifications using machine learning models, focusing on regression techniques.

Objectives:

- Understand the factors influencing laptop prices.
- Develop a robust predictive model for accurate price estimation.

DATASET OVERVIEW

Unnamed: 0	Company	TypeName	Inches	ScreenResolution	Cpu	Ram	Memory	Gpu	OpSys	Weight	Price
0	Apple	Ultrabook	13.3	IPS Panel Retina Display 2560x1600	Intel Core i5 2.3GHz	8GB	128GB SSD	Intel Iris Plus Graphics 640	macOS	1.37kg	71378.6832
1	Apple	Ultrabook	13.3	1440x900	Intel Core i5 1.8GHz	8GB	128GB Flash Storage	Intel HD Graphics 6000	macOS	1.34kg	47895.5232
2	HP	Notebook	15.6	Full HD 1920x1080	Intel Core i5 7200U 2.5GHz	8GB	256GB SSD	Intel HD Graphics 620	No OS	1.86kg	30636.0000
3	Apple	Ultrabook	15.4	IPS Panel Retina Display 2880x1800	Intel Core i7 2.7GHz	16GB	512GB SSD	AMD Radeon Pro 455	macOS	1.83kg	135195.3360
4	Apple	Ultrabook	13.3	IPS Panel Retina Display 2560x1600	Intel Core i5 3.1GHz	8GB	256GB SSD	Intel Iris Plus Graphics 650	macOS	1.37kg	96095.8080

df.shape

(1303, 12)

FEATURES:

Company, TypeName, Inches, ScreenResolution, Cpu, RAM, Memory, Gpu, Weight

Target Variable: Price

Size: [1303 rows x 12 columns]

SOURCE: <https://www.kaggle.com/code/enisigde/laptop-price-predict-linearregression>

DATA WRANGLING AND CLEANING



Steps Taken:

- Handled missing values and removed duplicates.
- Standardized units (e.g., RAM in GB, Weight in kg).
- Feature extraction: Calculated PPI and Touchscreen flags.
- Converted categorical variables using encoding techniques.

Outcome: Cleaned and ready dataset for analysis.

INSIGHTS FROM EXPLORATORY DATA ANALYSIS (EDA)

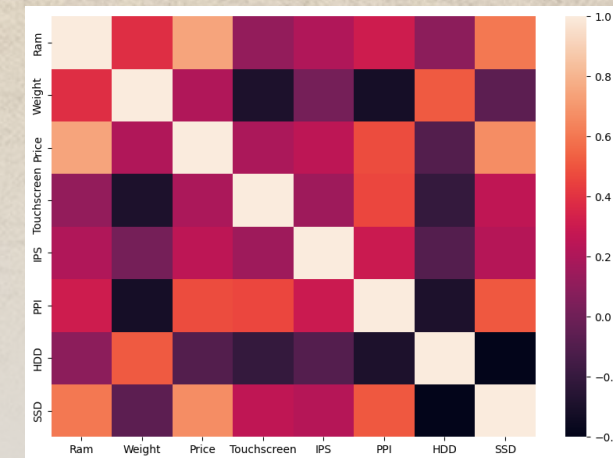
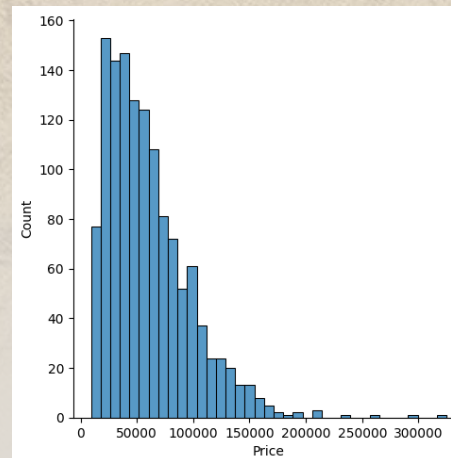
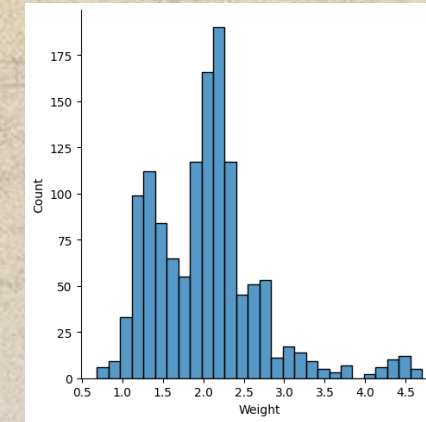
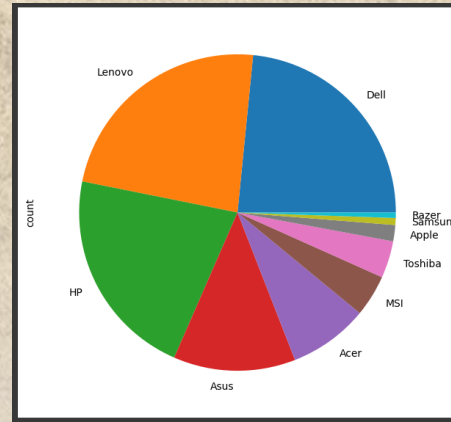


Visualizations:

- Distribution of Prices: Skewed, addressed using logarithmic transformation.
- Correlations: RAM, PPI, and SSD show strong positive relationships with Price.
- Company Insights: Apple laptops dominate the higher price range.

Key Findings:

- Price varies significantly with brand and specifications.
- Touchscreen and IPS displays influence pricing.



FEATURE ENGINEERING



Key Features:

- PPI (calculated from resolution and screen size).
- Binary flags for Touchscreen and IPS.
- Encoding categorical features: Company, TypeName, etc.

Purpose: Enhancing model performance by deriving informative features

Company	TypeName	Inches	ScreenResolution	Cpu	Ram	Memory	Gpu	OpSys	Weight	Price	Touchscreen	IPS	Xres	Yres
Apple	Ultrabook	13.3	IPS Panel Retina Display 2560x1600	Intel Core i5 2.3GHz	8	128GB SSD	Intel Iris Plus Graphics 640	macOS	1.37	71379	0	1	IPS Panel Retina Display 2560	1600
Apple	Ultrabook	13.3	1440x900	Intel Core i5 1.8GHz	8	128GB Flash Storage	Intel HD Graphics 6000	macOS	1.34	47896	0	0	1440	900
HP	Notebook	15.6	Full HD 1920x1080	Intel Core i5 7200U 2.5GHz	8	256GB SSD	Intel HD Graphics 620	No OS	1.86	30636	0	0	Full HD 1920	1080
Apple	Ultrabook	15.4	IPS Panel Retina Display 2880x1800	Intel Core i7 2.7GHz	16	512GB SSD	AMD Radeon Pro 455	macOS	1.83	135195	0	1	IPS Panel Retina Display 2880	1800
Apple	Ultrabook	13.3	IPS Panel Retina Display 2560x1600	Intel Core i5 3.1GHz	8	256GB SSD	Intel Iris Plus Graphics 650	macOS	1.37	96096	0	1	IPS Panel Retina Display 2560	1600

MODEL SELECTION AND TRAINING



Models Tested:

- Linear Regression, Decision Tree, Random Forest, Gradient Boosting, XGBoost.

Final Model: Random Forest Regressor

• Hyperparameters:

- `n_estimators`: 300
- `max_depth`: 20
- `max_features`: 0.75

- Training set R^2 : 0.92

- Testing set R^2 : 0.88

MODEL PERFORMANCE



Metrics:

- R^2 Score: 0.88 (Testing)
- Mean Absolute Error (MAE): 0.15 (log-transformed prices).

Interpretation:

- The model explains 88% of the variance in prices.
- MAE indicates low prediction errors.

CONCLUSION AND FUTURE WORK



Conclusion:

- Successfully predicted laptop prices with high accuracy.
- Random Forest proved to be the most effective model.

Future Work:

- Incorporate deep learning models for further accuracy.
- Include additional features like battery life and warranty.

The background of the slide features faint, sepia-toned sketches of two types of early aerial vehicles. On the left is a hot air balloon with a large, ornate, spherical envelope decorated with various patterns and a small basket hanging below. On the right is a rigid airship or blimp, characterized by its elongated, cigar-shaped hull, a series of vertical struts supporting a lower deck, and a single propeller at the rear. The entire scene is set against a textured, parchment-like background.

Thank you for your attention!

