Laptop Price Analysis

Laptop Price Analysis Project

By

Yashvi Vaghasiya

Introduction

- With the growth of e-commerce platforms, laptops are available in many configurations
- Laptop prices depend on multiple factors such as processor, RAM, storage, GPU, and display
- Manual price comparison becomes difficult for customers
- This project uses Machine Learning to analyze laptop specifications and predict prices
- The project demonstrates real-world data analysis and ML skills using Python

Dataset Description

- Publicly available laptop price dataset
- Total Records: 1275 laptops
- Total Features: 23 columns
- Dataset includes real laptop specifications and prices
- Original dataset contained compact text columns which were cleaned and restructured

Tools Used

- **Programming Language:** Python
- **Development Tool:** VS Code
- **Libraries Used:**
  - Pandas – Data manipulation
  - NumPy – Numerical computations
  - Matplotlib & Seaborn – Visualization
  - Scikit-learn – Machine Learning
- **Supporting Tools:** Excel (basic analysis)

Methodology

- Data Loading
- Data Cleaning and Preprocessing
- Exploratory Data Analysis (EDA)
- Feature Selection and Encoding
- Train-Test Split
- Model Training
- Model Evaluation
- Result Visualization

Data Preprocessing

- **Data Cleaning & Preprocessing**
- Removed duplicate records (28 duplicate rows)
- Verified no missing values in the dataset
- Converted categorical features into numerical form
- Removed high-cardinality columns (Product, CPU model, GPU model)
- Applied One-Hot Encoding for categorical variables
- Prepared final dataset for machine learning

Exploratory Data Analysis

- Analyzed distribution of laptop brands and operating systems
- Visualized RAM and storage distributions
- Studied touchscreen and IPS panel availability
- Compared CPU and GPU manufacturers
- Identified relationships between specifications and price

Machine Learning Model Used

- Algorithm Used: Linear Regression
- Reason for choosing Linear Regression:
  - Simple and interpretable
  - Suitable for regression problems
  - Good baseline model for price prediction
- Model trained on laptop specifications to predict price

Model Training & Evaluation

- Dataset split into:
  - 80% Training Data
  - 20% Testing Data
- Model performance evaluated using:
  - Mean Squared Error (MSE)
  - R-Squared (R^2) Score
- Predicted prices compared with actual prices

Results

- Laptop price increases with higher RAM and CPU frequency
- SSD storage significantly increases laptop price
- Touchscreen and Retina display laptops are more expensive
- Gaming laptops generally have higher prices
- Linear Regression provides reasonable prediction accuracy

Result Visualization:

- Scatter plot of Actual vs Predicted Prices
- Most data points align closely with the diagonal line
- Indicates good prediction performance
- Visualization helps validate model accuracy

Conclusion

- Laptop prices can be effectively predicted using machine learning
- Data preprocessing and feature engineering are crucial
- Linear Regression serves as a strong baseline model
- The project demonstrates an end-to-end ML workflow
- Strengthened practical skills in Data Analysis and Machine Learning

Reference

GitHub Link: https://github.com/Yashu-teach/Laptop-Price-Analysis-Project


THANK
YOU