

fingerTips

Machine Learning Evaluation

Predicting price of Laptops

Data Set Information and their description

Introduction:

The dataset at hand provides a comprehensive compilation of information concerning a diverse array of laptops, encompassing a spectrum of specifications and corresponding prices. Crafted with the aim of facilitating sales price prediction, this dataset offers invaluable insights into the intricate details of laptops, ranging from processor specifications to display attributes. With its rich repository of data, this dataset emerges as a pivotal resource for discerning trends within the laptop market and forecasting prices with precision.

Dataset Description:

1. Name: Name of the laptop model.
2. Brand: Brand of the laptop.
3. Price: Price of the laptop.
4. Rating: Rating of the laptop.

5. Processor_brand: Brand of the laptop's processor.
6. Processor_gen: Generation of the laptop's processor.
7. Core_per_processor: Number of cores per processor.
8. Energy_Efficient_Units: Indicates if the laptop has energy-efficient units.
9. RAM_GB: RAM capacity of the laptop in gigabytes.
10. RAM_type: Type of RAM.
11. Storage_capacity_GB: Storage capacity of the laptop in gigabytes.
12. Storage_type: Type of storage.
13. Graphics_brand: Brand of the laptop's graphics.
14. Display_size_inches: Size of the laptop's display in inches.
15. Horizontal_pixel: Number of horizontal pixels.
16. Vertical_pixel: Number of vertical pixels.
17. Touch_screen: Indicates if the laptop has a touch screen.
18. Operating_system: Operating system of the laptop.

Objective:

The primary objective of this dataset is to enable the analysis and prediction of laptop sales prices through the exploration of various features and attributes associated with these devices. By leveraging the dataset's extensive information on laptop specifications, including processor details, display specifications, and other relevant factors, the goal is to develop predictive models capable of accurately estimating the prices of laptops. Through this endeavor, the dataset seeks to enhance understanding of market trends and contribute to informed decision-making processes within the realm of laptop sales and pricing.

Questions:

1. How can the dataset be imported into .ipynb code using the pandas library?
2. Considering the nature of the dataset, would it be more suitable to employ Supervised or Unsupervised Machine Learning techniques?
3. If Supervised learning is deemed appropriate, what type of problem statement does it represent - Regression or Classification?

4. What methods can be employed to assess the quality of the data, including its shape and data types?
5. How can null values within the dataset be identified and handled effectively?
6. Appropriate form of analysis should be presented and described in form of comments headings and analysis?
7. What techniques can be utilized to detect and eliminate any outliers present in the dataset?
8. In cases where categorical variables are present, how can they be encoded using appropriate techniques?
9. When necessary, how can the data be scaled using appropriate scaling techniques?
10. How can the dataset be divided into predictors and target variables?
11. What is the recommended approach for dividing the dataset into training and testing sets?
12. Which machine learning model is most suitable for training the data, considering its characteristics?
13. How can the accuracy of the trained model be evaluated to determine its effectiveness?
14. Proper approach to enhance the model by certain techniques must be done?

Certainly, here are the evaluation criteria:

1. Data Cleaning - 20 marks
2. Exploratory Data Analysis - 20 marks
3. Data Preprocessing - 10 marks
4. Data Splitting - 10 marks
5. Model Implementation - 10 marks
6. Model Evaluation and Analysis - 10 marks
7. Application of Enhancement Techniques in Modelling - 10 marks
8. Code Structure, Comments, and Output Clarity - 10 marks