Laptop Price Prediction and Clustering Using Web Scraping and Machine Learning

Name: M. Sudharsanan

# 1. Web Scraping Laptop Data from Flipkart

Laptop details including specifications and prices were scraped from Flipkart using Selenium and BeautifulSoup to create a structured dataset for analysis and prediction.

# 2. Dataset

Columns included in the dataset:  
- Title  
- Original Price  
- Discount Price  
- Rating  
- Number of Reviews  
- Offers

# 3. Data Preprocessing

Steps taken:  
- Standardize data types  
- Handle missing values  
- Remove duplicates  
- Outlier detection and removal

# 4. Visualization

Note: Visualizations from the dataset such as price distributions, rating comparisons, etc.

# 5. Unsupervised Machine Learning

K-Means Clustering was used to group laptops into clusters.  
  
Metrics:  
- Silhouette Score  
  
Visualization:  
- Clusters visualized using K = 2

# 6. Supervised Machine Learning

Multiple models were trained:  
- Logistic Regression  
- K-Nearest Neighbors (KNN)  
- Support Vector Machine (SVM)  
- Decision Tree  
- Random Forest  
- XGBoost Classifier

# 7. Hyperparameter Tuning

Initial SVM model achieved 96% accuracy before hyperparameter optimization.  
GridSearchCV was used to find the best hyperparameters.  
Best parameters:  
- C = 1  
- gamma = 1  
- kernel = 'linear'  
  
Final model achieved 100% accuracy with improved precision, recall, and F1-score.

# 8. Conclusion

Hyperparameter tuning using GridSearchCV enabled the SVM model to achieve 100% accuracy, demonstrating its effectiveness when optimally configured.