

Shri Ramdeobaba College of Engineering & Management, Nagpur.
Department of Computer Science &
Engineering Session: 2024-25 [Odd Semester]
VII Semester [Shift-I]

Project Title: Laptop Price Predictor

Problem Definition:

The Laptop Price Predictor is a machine learning-based project designed to predict the prices of laptops based on their technical specifications, such as RAM, processor type, GPU, screen resolution, storage, and other features. The project aims to build a model that accurately estimates the laptop's price using linear regression and random forest algorithms. The project also includes data preprocessing, feature engineering, and the deployment of the model for real-time predictions.

Project Objectives:

1. **Data Preprocessing:** Clean and preprocess the dataset by handling missing values, removing duplicates, and transforming the data into suitable formats.
2. **Feature Engineering:** Extract meaningful features such as touchscreen, IPS display, pixel density (PPI), and categorize processors, operating systems, and memory configurations.
3. **Model Development:** Implement machine learning algorithms (Linear Regression and Random Forest) to predict laptop prices and evaluate their performance using metrics like R2 score and Mean Absolute Error (MAE).
4. **Optimization:** Fine-tune the machine learning model to improve prediction accuracy by experimenting with different hyperparameters and algorithms.
5. **Model Deployment:** Save the trained model using Python's pickle library to be deployed for real-time predictions.
6. **Visualization and Analysis:** Create visualizations to analyze trends in the data, such as the distribution of brands, GPU types, and RAM capacity.
7. **Testing and Validation:** Split the dataset into training and testing sets to evaluate the model's performance and ensure it generalizes well on unseen data.

Methodology:

1. Data Collection:

- The dataset is loaded from a CSV file containing laptop specifications and corresponding prices.

2. Data Preprocessing:

- Handle missing and duplicate values.
- Remove unwanted columns and convert data types (e.g., 'RAM' from string to integer, 'Weight' from string to float).
- Extract features from columns like 'ScreenResolution' and categorize operating systems and CPUs

3. Feature Engineering:

- Extract new features such as 'Touchscreen', 'IPS', and 'PPI' from the 'ScreenResolution' column.
- Process the 'Memory' column to split HDD, SSD, Hybrid, and Flash storage.
- Group the CPU and GPU brands into relevant categories.

4. Model Building:

- Split the data into training and testing sets using `train_test_split`.
- Implement a ColumnTransformer to encode categorical variables.
- Create pipelines to preprocess the data and fit the machine learning models (Linear Regression and Random Forest).

5. Model Evaluation:

- Evaluate the model's performance using the R2 score and Mean Absolute Error (MAE).
- Test different models and algorithms, including Linear Regression and Random Forest, to determine which provides the best results.

6. Model Deployment:

- Save the model using Python's `pickle` library to allow it to be reused for predictions without retraining.

7. Model Prediction:

- Test the model on unseen data to make price predictions and compare with actual laptop prices.

Technology Stack:

1. **Programming Language:**
 - Python (for model development and data processing)
2. **Libraries & Frameworks:**
 - **Pandas**: For data loading, preprocessing, and manipulation.
 - **NumPy**: For numerical operations.
 - **Matplotlib**: For data visualization.
 - **Scikit-learn**: For implementing machine learning models, pipelines, and metrics.
 - **Pickle**: For saving and loading machine learning models.
3. **Algorithms:**
 - Linear Regression
 - Random Forest Regression
4. **Data Handling:**
 - CSV file for the laptop dataset.
5. **Deployment:**
 - **Google Colab**: For training and running the model.
 - **Pickle**: For model serialization and deployment.

| Roll No. | Name of Students |
|----------|--------------------|
| 30 | Adwait Puntambekar |
| 32 | Aman Joharapurkar |
| 38 | Atharva Telrandhe |
| 57 | Samrat Roy |