Laptop Price Prediction Project

NAME: Said Mohamed Ibrahim

ID: 41810238

NAME: omar nasr eldin

ID: 41910045

Dataset: Kaggle: Laptops Price Dataset

1. Problem Definition

The goal of this project is to predict the price of laptops based on their specifications using a machine learning model. Given the wide variety of laptops with differing specifications and configurations, it becomes essential to identify patterns and relationships between these specifications and their impact on price. This analysis can help consumers make informed decisions and manufacturers to strategically price their products.

2. Method

To achieve the goal, the following steps were taken:

- **Data Collection**: The dataset was obtained from Kaggle and contains detailed information on various laptops, including fields like Laptop Name, Brand, Model, CPU, GPU, RAM, Storage, and Final Price.

- **Data Cleaning**: Missing values were handled by filling them with median values where applicable. Columns were renamed and redundant fields were dropped.
- **Data Preprocessing**: Numerical features were scaled, and categorical features were encoded using OneHotEncoder to ensure they were in a suitable format for machine learning algorithms.
- **Model Training**: A Random Forest Regressor was used to train the model. The dataset was split into training and test sets, with the training data being used to fit the model.
- **Hyperparameter Tuning**: GridSearchCV was employed to find the best hyperparameters for the Random Forest model. The search identified the optimal parameters: {'max_depth': 20, 'n_estimators': 200}.

3. Experiment

- **Model Training**: The Random Forest model was trained using the preprocessed training data with the optimal hyperparameters found through GridSearchCV.
- **Evaluation**: The model was evaluated on the test data using metrics like Mean Squared Error (MSE) and R^2 Score. The model achieved a Mean Squared Error of 0.005272846665155251 and an R^2 Score of 0.79422637316632, indicating a strong predictive performance.
- **Results**: The model effectively captures the relationship between laptop specifications and their prices, making it a useful tool for price prediction and analysis.

4. References

- Kaggle: Laptops Price Dataset
- Pedregosa, F., et al. (2011). "Scikit-learn: Machine Learning in Python." Journal of Machine Learning Research, 12, 2825-2830.