**Executive Summary: Laptop Price Prediction Project**

1. **Introduction**

The Laptop Price Prediction project aims to develop a machine learning model capable of accurately predicting laptop prices based on various features such as company, type, screen resolution, CPU, RAM, memory type, GPU, operating system, weight, and additional specifications like touchscreen functionality and pixel density. This project follows a structured approach involving data preprocessing, exploratory data analysis (EDA), feature engineering, model selection, and evaluation.

1. **Data Acquisition and Preprocessing**

The dataset consists of 1,303 laptop records with 12 attributes. Initial preprocessing steps included:

* Removing unnecessary columns (e.g., unnamed index column).
* Converting categorical data into numerical formats.
* Extracting meaningful features such as pixel density (PPI) from screen resolution.
* Standardizing RAM and weight units.
* Splitting and categorizing memory types (SSD, HDD, Hybrid, Flash Storage).
* Encoding categorical features like CPU brand, GPU brand, and operating system.

1. **Exploratory Data Analysis (EDA)**

EDA was conducted to understand feature distributions, detect outliers, and identify correlations. Key insights included:

* RAM and SSD storage showed a strong positive correlation with price.
* HDD storage had a negative correlation with price.
* Higher PPI and touchscreen functionality generally contributed to higher prices.
* Certain brands (Apple, Dell, HP) had higher average prices compared to others.

1. **Feature Engineering**

New features were engineered to improve model performance, including:

* Pixel density (PPI) as a derived feature from screen resolution.
* Binary indicators for touchscreen and IPS display.
* Categorization of CPU brands (Intel Core i3/i5/i7, AMD, Other Intel processors).
* GPU brand extraction and classification.
* Categorization of operating systems into Windows, Mac, and Others.

1. **Model Selection and Evaluation**

Several machine learning models were trained and evaluated using appropriate metrics (R2 Score, Mean Absolute Error (MAE)):

* **Linear Regression**: R2 Score = 0.80, MAE = 0.2115
* **Ridge Regression**: R2 Score = 0.81, MAE = 0.2098
* **Lasso Regression**: R2 Score = 0.80, MAE = 0.2116
* **K-Nearest Neighbors (KNN)**: R2 Score = 0.83, MAE = 0.1707
* **Decision Tree Regression**: R2 Score = 0.88, MAE = 0.1620
* **Random Forest Regression**: R2 Score = 0.96, MAE = 0.0597
* **Gradient Boosting**: R2 Score = 0.96, MAE = 0.0688
* **XGBoost**: R2 Score = 0.96, MAE = 0.0672
* **Voting Regressor** (Best Performing Model): R2 Score = 0.97, MAE = 0.0547

Hyperparameter tuning was applied to Decision Trees, Random Forest, and AdaBoost using GridSearchCV, significantly improving performance. The final model, a Voting Regressor combining Random Forest, Gradient Boosting, and XGBoost, achieved the highest accuracy.

1. **Conclusion**

The project successfully built a robust predictive model for laptop prices. The best-performing model, the Voting Regressor, provides a high R2 score (0.97) and low MAE (0.0547), indicating strong predictive power. This model can assist consumers, retailers, and manufacturers in estimating laptop prices based on specifications. Future enhancements may include expanding the dataset, incorporating real-time pricing data, and exploring deep learning techniques for further improvements.