Terminal Output :

----- Q3: Single Decision Tree -----
Single Tree Prediction: Yes

----- Q4: Bagging with 10 trees -----
Bagging Prediction: Yes
Bagging OOB Error: 0.5000

----- Q4: Random Forest with 2 random predictors -----
Random Forest Prediction: Yes
Random Forest OOB Error: 0.1250

1. Single Decision Tree (DT):

- **Prediction**: The **single decision tree** predicts **"Yes"** for the given input sample.
- **Weaknesses**:
    - A **single decision tree** is prone to **overfitting** or **underfitting** depending on its depth and the complexity of the data.

2. Bagging with 10 Trees:

- **Prediction**: The **bagging** ensemble predicts **"Yes"**, just like the single decision tree.
- **Out-of-Bag (OOB) Error**: **0.5000**.
    - An **OOB error of 0.5000** means that **50%** of the time, the bagging model made an incorrect prediction when tested on the OOB samples.Bagging reduces **variance** and generally improves stability compared to a single decision tree, but it still can struggle if the base models (the individual trees) are weak or the data is very noisy.

3. Random Forest with 2 Random Predictors:

- **Prediction**: The **random forest** model also predicts **"Yes"** for the same input sample.
- **Out-of-Bag (OOB) Error**: **0.1250**.
    - The **OOB error** here is much lower than for bagging, at **12.5%**. This suggests that **random forest** is performing significantly better on this dataset.
- **Strengths/Weaknesses**:
    - **Random Forests** tend to work better than bagging or individual decision trees because of their ability to avoid overfitting and to reduce bias through randomness in both data samples and features. It balances **variance** and **bias** better than bagging or individual trees, leading to better performance, especially for more complex or high-dimensional data.

Q-5
Terminal Output :

Train & Test MSE for each degree:
Degree 1: Train MSE = 0.194403, Test MSE = 0.211439
Degree 2: Train MSE = 0.195395, Test MSE = 0.201545
Degree 3: Train MSE = 0.011915, Test MSE = 0.013481
Degree 4: Train MSE = 0.011856, Test MSE = 0.013973

Best degree: 3

Interpretation:

Best Degree: 3: Based on the cross-validation results, the cubic model (degree 3) provides the best balance between low training error and low test error. The test MSE for degree 3 is very close to that of degree 4 but with a slightly better generalization. Higher-degree models (degree 4) show marginal improvement in the training error but with a slightly worse test error,

suggesting they may be overfitting the data.