Part-I: Naive Bayes Classifier

1.What is the core assumption of Naive Bayes?

- Core Assumptions of Naïve Bayes is that:
- All features are conditionally independent given the class label.

2. Differentiate between GaussianNB, MultinomialNB, and BernoulliNB:

- **GaussianNB** assumes the features follow a normal (Gaussian) distribution.
    - **MultinomialNB** is used for count-based features (like word frequencies) and is common in text classification.
- **BernoulliNB** works with binary/boolean features (e.g., word presence/absence in document.

3. Why is Naive Bayes considered suitable for high-dimensional data?

- It makes strong independence assumptions.
- Each feature is treated as an independent features. This avoids dimensionality and simplifies training the dataset.

Part-II: Decision Trees

1. What is entropy and information gain?

- Entropy measures the impurity or disorder in a dataset.
- Information gain is the reduction in entropy after a dataset is split based on a feature; it's used to choose the best feature for splitting in a decision tree.

2. Explain the difference between Gini Index and Entropy:

- **Gini Index** : It measures the probability of misclassifying a randomly chosen element. It is faster.
- **Entropy:** It uses logarithms to measure impurity.

3. How can a decision tree overfit? How can this be avoided?

A decision tree can overfit by growing too deep and capturing noise in the training data. This can be avoided by pruning, limiting max depth, or setting minimum samples for splits or leaf nodes.

Part-III: Ensemble Learning – Bagging, Boosting, Random Forest

1. What is the difference between Bagging and Boosting?

- **Bagging** builds multiple models independently on **random subsets** of the data and combines their outputs (e.g., Random Forest).
- **Boosting** builds models **sequentially**, where each new model tries to correct the errors of the previous one (e.g., AdaBoost, Gradient Boosting).

2. How does Random Forest reduce variance?

Random Forest reduces variance by averaging the predictions of many uncorrelated decision trees built on bootstrapped data and random feature subsets, which stabilizes the final output and prevents overfitting.

3. What is the weakness of boosting-based methods?

Boosting methods can be sensitive to noisy data and outliers, and they are more prone to overfitting if not properly tuned (e.g., too many trees or too deep). They also tend to be computationally more expensive than bagging methods.
