Question 1:

1. Linear Regression:

Strengths:

1. Simple and easy to interpret.
2. Works well when the relationship between features and target is linear.

Limitations:

1. Assumes linearity, which may not hold for complex data.
2. Sensitive to outliers.

Application Scenario:

Best for predicting continuous variables in a linear relationship (e.g., predicting house prices based on size). Would not use for classification problems.

2. Support Vector Machines (SVM):

Strengths:

1. Effective in high-dimensional spaces.
2. Works well with both linear and non-linear data (with kernel trick).

Limitations:

1. Computationally expensive for large datasets.
2. Difficult to interpret results.

Application Scenario:

Ideal for image classification where data is complex and high-dimensional. Less suitable for simple, linear datasets.

3. k-Nearest Neighbors (k-NN):

Strengths:

1. Simple and intuitive.
2. No training phase; works well with small datasets.

Limitations:

1. Computationally expensive for large datasets (due to distance calculation).
2. Sensitive to noise and irrelevant features.

Application Scenario:

Suitable for recommendation systems or anomaly detection in small datasets. Would not use in large-scale or high-dimensional data.

4. Decision Trees:

Strengths:

1. Easy to visualize and interpret.
2. Can handle both categorical and numerical data.

Limitations:

1. Prone to overfitting, especially in complex trees.
2. Unstable, as small changes in data can result in different trees.

Application Scenario:

Ideal for medical diagnosis decision-making but not suitable for continuous output like regression without proper regularization.

Question 2:

1. Bias and Variance Definitions:

Bias: Error due to overly simplistic assumptions in the model, leading to underfitting.

Variance: Error due to model sensitivity to small fluctuations in the training data, leading to overfitting.

2. Influence on Model Performance:

- High bias can cause the model to miss important patterns, resulting in underfitting.
- High variance means the model captures noise in the data, leading to overfitting.

3. Impact of Model Complexity:

- Simple models (like linear regression) tend to have high bias and low variance.
- Complex models (like deep decision trees) tend to have low bias but high variance.

4. Strategies to Balance Bias and Variance:

- Cross-validation helps in evaluating model performance across different subsets.
- Regularization (e.g., L1, L2 penalties) can help reduce variance in complex models.

- Ensemble methods like bagging and boosting reduce variance by averaging predictions across multiple models.