Answer a)

Linear Regression:-

Strengths:

i. Good interpretability: Coefficients provide clear insights into feature importance.
ii. Computationally efficient: Training is fast and prediction are quick.

Limitations:

i. Linearity Assumption: Based on assumption of linear relationship between features and target which may not always be true.
ii. Sensitivity to outliers: Outliers in training data can significantly impact the model performance.

Suitable application scenario: Where linear relationship exists, and high interpretability & speed are desired. Eg: Predicting height from age, predicting house price based on no. of rooms, size of house in square foot etc.

Support Vector Machines (SVM):-

Strengths:

i. Effective for noisy/high-dimensional dataset: Works well on relatively large number of features while remaining robust towards noise.
ii. Kernel tricks: Kernel tricks enable SVMs to deal with linearly not separable classes while controlling overfitting.

Limitations:

i. Computational complexity and cost: SVMs can be very expensive in time and space for large datasets.
ii. Hard to interpret: Can be hard to interpret and understand, especially when non-linear kernels are involved.

Suitable application scenario: Where data is high-dimensional/noisy, linearly not separable. Eg: Image classification/categorization.

K Nearest Neighbour (k-NN):-

Strengths:

i. Simplicity and intuitiveness: Easy to understand and implement.
ii. Non-parametric: Can model complex patterns without making any assumptions about the underlying data distribution.

Limitations:

i. Computationally expensive: As k-NN needs to store the entire training dataset and compute distances for every new prediction, it is very slow and consumes large compute resources especially for large datasets.
ii. Sensitivity to noise in features/target classes: k-NN struggles to perform on noisy dataset if the chosen value of ‘k’ or distance metric are not optimal.

Suitable application scenario: Where similarity-based recommendations are needed without explicit training. Eg: Product recommendations based on user preferences.

Decision Trees:-

Strengths:

i. Simplicity and interpretability: DTs provide clear and easy to interpret visual representation and are very fast at prediction time.
ii. Ability to handle different types of features: DTs work with numerical as well as categorical data and can be even more powerful when combined via ensemble methods.

Limitations:

i. Prone to overfitting: DTs are prone to overfitting, especially if they are not pruned or dataset is too small/noisy.
ii. Learning intractability: Learning optimal DT is intractable (NP-hard).

Suitable application scenario: Where high interpretability and non-linear decision making are desired. Eg: Credit risk assessment.

Answer b)

The bias-variance tradeoff reflects the balance between underfitting and overfitting in supervised learning. While Bias is a measure of accuracy of the prediction, Variance is a measure of variability of the prediction across datasets. High bias results from overly simplistic assumptions during training and leads to underfitting. High variance is a result of the model capturing noise along with underlying patterns in the training data and leads to overfitting/poor generalization. Complex models tend to have low bias but high variance. On the other hand, simple models have lower variance but high bias. Some of the strategies to balance bias and variance are - cross-validation, regularization, early stopping, ensemble methods (boosting, bagging) and hyperparameter tuning.