

1. **KNN Algorithm:**
  - KNN (K-Nearest Neighbors) is a supervised machine learning algorithm used for both classification and regression tasks.
  - It makes predictions based on the similarity of a new data point to its k nearest neighbors in the training dataset.
2. **Choosing the Value of K:**
  - The value of K is typically chosen through hyperparameter tuning, using techniques like cross-validation.
  - A small value of K may lead to noisy predictions and overfitting, while a large value of K may result in overly smoothed decision boundaries and underfitting.
3. **Difference between KNN Classifier and Regressor:**
  - KNN Classifier predicts the class label of a data point based on the majority class of its k nearest neighbors.
  - KNN Regressor predicts the continuous value of a data point based on the average or weighted average of the target values of its k nearest neighbors.
4. **Measuring Performance of KNN:**
  - Performance of KNN can be measured using classification accuracy, precision, recall, F1 score, or regression metrics like mean squared error (MSE) or R-squared.
5. **Curse of Dimensionality:**
  - The curse of dimensionality refers to the phenomenon where the feature space becomes increasingly sparse as the number of dimensions (features) grows.
  - In KNN, as the dimensionality of the feature space increases, the distance between data points becomes less meaningful, making it difficult to find meaningful nearest neighbors.
6. **Handling Missing Values in KNN:**
  - One approach is to impute missing values using techniques like mean imputation, median imputation, or KNN imputation itself.
  - Alternatively, missing values can be handled by treating them as a separate category or using advanced imputation techniques like iterative imputation.
7. **Comparison of KNN Classifier and Regressor:**
  - KNN Classifier is suitable for classification problems where the decision boundaries are relatively smooth and the classes are well-separated.
  - KNN Regressor is suitable for regression problems where the relationships between features and target values are not linear and have complex patterns.
8. **Strengths and Weaknesses of KNN:**
  - Strengths: Simple to implement, non-parametric, handles multi-class problems, robust to noisy training data.
  - Weaknesses: Computationally expensive during prediction, sensitive to the choice of distance metric and value of K, requires a large amount of memory to store the entire training dataset.
9. **Difference between Euclidean and Manhattan Distance:**
  - Euclidean distance measures the shortest straight-line distance between two points in Euclidean space.
  - Manhattan distance (also known as city block or taxicab distance) measures the distance between two points as the sum of the absolute differences of their coordinates.
10. **Role of Feature Scaling in KNN:**
  - Feature scaling is important in KNN because it helps to ensure that all features contribute equally to the distance calculation.

- Without scaling, features with larger magnitudes may dominate the distance calculation, leading to biased results.
- Common scaling techniques include standardization (subtracting the mean and dividing by the standard deviation) and normalization (scaling features to a fixed range, typically  $[0, 1]$ ).