20th April quiz   7 out of 7 correct

1. What does KNN stand for in the context of machine learning?

○   Kernelized Nearest Neighbors

○   K-means Nearest Neighbors

⊙   **K-nearest Neighbors**

○   K-weighted Nearest Neighbors

Explanation: KNN is a machine learning algorithm that works on the principle of finding the k nearest data points in the training set to a given input, and using their labels or values to make predictions.

2. KNN can be used for:

○   Classification

○   Regression

⊙   **Both Classification and Regression**

○   Neither Classification nor Regression

Explanation: KNN can be used for both classification and regression tasks. In classification, KNN predicts the class of a new data point based on the majority class of its k nearest neighbors in the training set. In regression, KNN predicts the value of a new data point based on the average of the values of its k nearest neighbors in the training set.

3. The value of k in KNN:

○   Must always be odd

○   Must always be even

○ **Can be any positive integer**

○ Must be less than the size of the training set

Explanation: The value of k in KNN is a hyperparameter that can be tuned to optimize the performance of the model. It can be any positive integer, and does not necessarily have to be odd or even. However, choosing an appropriate value of k is important to avoid underfitting or overfitting the model.

4. Which of the following is not a distance metric commonly used in KNN?

○ Euclidean distance

○ Manhattan distance

○ Cosine distance

○ **Pearson correlation**

Explanation: Pearson correlation is a similarity metric commonly used in collaborative filtering and other recommendation systems, but not in KNN. The distance metrics commonly used in KNN include Euclidean distance, Manhattan distance, and cosine distance.

5. In KNN, the decision boundary between classes is:

○ Linear

○ **Non-linear**

○ Always a circle

○ Always a straight line

Explanation: In KNN, the decision boundary between classes is non-linear, because the algorithm uses the distances between data points to determine their similarity, rather than assuming a linear relationship between the features.

6. The main disadvantage of KNN compared to other machine learning algorithms is:

○ It is computationally expensive at inference time

○ It cannot handle missing values in the data

○ It is prone to overfitting the training set

○ It requires a large amount of training data

Explanation: The main disadvantage of KNN compared to other machine learning algorithms is that it is computationally expensive at inference time, because it requires calculating the distances between the new data point and all the data points in the training set. This can make it slow and impractical for large datasets.

7. Which of the following is a possible way to reduce the computational cost of KNN at inference time?

○ Use a larger value of k

● Use a smaller value of k

○ Use a faster distance metric

○ None of the above

Explanation: One way to reduce the computational cost of KNN at inference time is to use a smaller value of k, which reduces the number of data points that need to be considered. However,

Submit