GATE

CRASH COURSE GATE 2025

Machine Learning

KNN

Q1 In k-NN, what is the effect of using weighted voting for classification?
(A) It gives equal importance to all neighbors
(B) It assigns more importance to closer neighbors
(C) It reduces the effect of outliers
(D) It increases the model's bias

Q2 What can be a consequence of using k-NN on a dataset with imbalanced classes?
(A) Increased accuracy
(B) Decreased accuracy
(C) Balanced predictions
(D) Improved precision

Q3 Which of the following scenarios is least suitable for k-NN?
(A) Small datasets
(B) Large datasets
(C) High-dimensional datasets
(D) Datasets with continuous target variables

Q4 What is the primary challenge of using k-NN in high-dimensional spaces?
(A) Overfitting
(B) Underfitting
(C) Curse of dimensionality
(D) Scalability

Q5 In k-NN regression, how is the final prediction made?
(A) By taking the most frequent label
(B) By taking the mean of the nearest neighbors' target values
(C) By taking the median of the nearest neighbors' target values
(D) By taking the mode of the nearest neighbors' target values

Q6 Which technique can be used to reduce the dimensionality of the dataset before applying k-NN?
(A) k-means clustering
(B) PCA (Principal Component Analysis)
(C) Linear Regression
(D) Decision Trees

Q7 What is the primary disadvantage of using the Manhattan distance in k-NN?
(A) It is computationally expensive
(B) It is sensitive to outliers
(C) It can only be used for continuous data
(D) It does not consider diagonal distance

Q8 How can the performance of a k-NN classifier be evaluated?
(A) By checking the training accuracy
(B) By checking the computational time
(C) By using a confusion matrix and calculating metrics such as precision, recall, and F1-score
(D) By the number of neighbors

GATE

Answer Key

Q1	(B)	Q5	(B)
Q2	(B)	Q6	(B)
Q3	(C)	Q7	(D)
Q4	(C)	Q8	(C)

GATE

Hints & Solutions

Q1 **Text Solution:**
Weighted voting gives more importance to closer neighbors by assigning them higher weights, potentially improving prediction accuracy.

Q2 **Text Solution:**
In imbalanced datasets, k-NN may be biased towards the majority class, leading to decreased accuracy for the minority class.

Q3 **Text Solution:**
k-NN performs poorly in high-dimensional spaces due to the curse of dimensionality, which makes distance measurements less meaningful.

Q4 **Text Solution:**
The curse of dimensionality refers to various phenomena that arise when analyzing data in high-dimensional spaces, where distances between points become less distinguishable.

Q5 **Text Solution:**
In k-NN regression, the predicted value is typically the mean of the target values of the k nearest neighbors.

Q6 **Text Solution:**
PCA is a dimensionality reduction technique that can reduce the number of features in a dataset, potentially improving k-NN performance.

Q7 **Text Solution:**
The Manhattan distance (L1 distance) measures distance along axes at right angles and does not account for diagonal movement, which can be less intuitive for some applications.

Q8 **Text Solution:**
Evaluating a classifier's performance typically involves using a confusion matrix to calculate precision, recall, F1-score