

19CAT703-Machine Learning

IAE2

Email *

manikandannagaraj748@gmail.com

k-NN algorithm does more computation on test time rather than train time

1 point

- True
- False

Which of the following distance metric can not be used in k-NN?

1 point

- A) Manhattan
- B) Minkowski
- C) Tanimoto
- D) All the above

Which of the following statement is true about k-NN algorithm?k-NN performs much better if all of the data have the same scalek-NN works well with a small number of input variables (p), but struggles when the number of inputs is very largek-NN makes no assumptions about the functional form of the problem being solved 2 points

- A) 1 and 2
- B) 1 and 3
- C) Only 1
- D) All of the above

What is the minimum no. of variables/ features required to perform clustering? 2 points

- A : 0
- B : 1
- C : 2
- D : 3

The goal of clustering a set of data is to 1 point

- A : divide them into groups of data that are near each other
- B : choose the best data from the set
- C : predict the class of data
- D : determine the nearest neighbors of each of the data

_____ is a clustering procedure where all objects start out in one giant cluster. Clusters are formed by dividing this cluster into smaller and smaller clusters. 1 point

- A : Non-hierarchical clustering
- B : Divisive clustering
- C : Agglomerative clustering
- D : K-means clustering

Which of the following clustering algorithms suffers from the problem of convergence at local optima?1. K- Means clustering algorithm2. Agglomerative clustering algorithm3. Expectation-Maximization clustering algorithm4. Diverse clustering algorithm 2 points

- A : 1 only
- B : 2 and 3
- C : 2 and 4
- D : 1 and 3

The k-means algorithm... 1 point

- A : always converges to a clustering that minimizes the mean-square vector-representative distance
- B : can converge to different final clustering, depending on initial choice of representatives
- C : is typically done by hand, using paper and pencil
- D : should only be attempted by trained professionals

what is true about single linkage hierarchical clustering

2 points

- A : we merge in each step the two clusters, whose two closest members have the smallest distance.
- B : we merge in the members of the clusters in each step, which provide the smallest maximum pairwise distance.
- C : the distance between two clusters is defined as the average distance between each point in one cluster to every point in the other cluster.
- D : none of the above

Agglomerative clustering is

1 point

- A : the initial state is a single cluster with all samples and the process proceeds by splitting the intermediate cluster until all elements are separated
- B : process starts from the bottom (each initial cluster is made up of a single element) and proceeds by merging the clusters until a stop criterion is reached.
- C : requires prior knowledge of no. of clusters you want to divide your data into.
- D : None of the above

Frequent visitors to some book sites often see lists of suggested titles based on their previous purchases at the site. Websites making book recommendations may be using all the following algorithms EXCEPT 1 point

- A : collaborative filtering
- B : instrumental filtering
- C : content-based filtering
- D : rule-based filtering

Which of the following clustering requires merging approach? 1 point

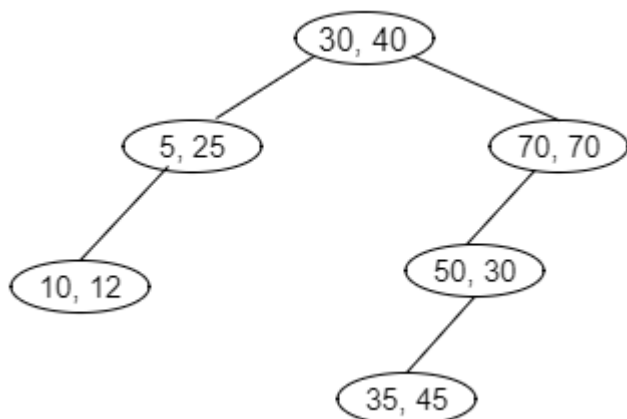
- a) Partitional
- b) Hierarchical
- c) Naive Bayes
- d) None of the mentioned

Which of the following is the simplest data structure that supports range searching? 1 point

- a) Heaps
- b) binary search trees
- c) AA-trees
- d) K-d trees

What will be the correct sequence of insertion for the following k-d tree?

2 points



- a) (30,40),(5,25),(70,70),(10,12),(50,30),(35,45)
- b) (40,30),(5,25),(12,10),(70,70),(30,50),(45,35)
- c) (30,40),(5,25),(10,12),(70,70),(50,30),(35,45)
- d) (40,30),(25,5),(12,10),(70,70),(50,30),(45,35)

Reducing search space by eliminating irrelevant trees is known as?

1 point

- a) pruning
- b) partial results
- c) freeing space
- d) traversing

The 2d search tree has the simple property that branching on odd levels is done with respect to the first key. 1 point

- a) True
- b) False

In what time can a 2-d tree be constructed? 2 points

- $O(N)$
- $O(N \log N)$
- $O(N^2)$
- $O(M \log N)$

In a k-d tree, k originally meant? 1 point

- number of dimensions
- size of tree
- length of node
- weight of node

Which of the following is a disadvantage of non-parametric machine learning algorithms?

1 point

- a) Capable of fitting a large number of functional forms (Flexibility)
- b) Very fast to learn (Speed)
- c) More of a risk to overfit the training data (Overfitting)
- d) They do not require much training data

Suppose we have a regularized linear regression model. What is the effect of increasing λ on bias and variance?

1 point

- a) Increases bias, increases variance
- b) Increases bias, decreases variance
- c) Decreases bias, increases variance
- d) Decreases bias, decreases variance

What metric can be used to find an optimal number of clusters ?

1 point

- R Squared
- MSE
- WCSS

In Python, what is the recommended init parameter to input ?

1 point

- random
- k-means++
- inertia
- boost

Sentiment Analysis is an example of:

1 point

- Regression
- Classification
- Reinforcement Learning
- All the above

Which of the following are ML methods

1 point

- Based on human supervision
- supervised learning
- unsupervised learning
- reinforcement learning

The goal of clustering analysis is to:

0 points

- a) Maximize the inter-cluster similarity
- b) Maximize the intra-cluster similarity
- c) Maximize the number of clusters
- d) Minimize the intra-cluster similarity

Random Forest has _____ as base learning models

1 point

- A) multiple decision trees
- B) bagging
- C) Entropy
- D) none of these

_____ helps improve machine learning results by combining several models.

1 point

- A) Machine Learning
- B) bagging
- C) Entropy
- D) Ensemble learning

Which of the following algorithm doesn't uses learning Rate as of one of its hyperparameter?

2 points

- Gradient Boosting
- Extra Trees
- AdaBoost
- Random Forest

In random forest or gradient boosting algorithms, features can be of any type. For example, it can be a continuous feature or a categorical feature. Which of the following option is true when you consider these types of features?

1 point

- A) Only Random forest algorithm handles real valued attributes by discretizing them
- B) Only Gradient boosting algorithm handles real valued attributes by discretizing them
- C) Both algorithms can handle real valued attributes by discretizing them
- D) None of these

Which of the following is true about the Gradient Boosting trees? In each stage, introduce a new regression tree to compensate the shortcomings of existing model. We can use gradient decent method for minimize the loss function

2 points

- A) 1
- B) 2
- C) 1 and 2
- D) None of these

Consider the hyperparameter “number of trees” and arrange the options in terms of time taken by each hyperparameter for building the Gradient Boosting model? 2 points
Note: remaining hyperparameters are same
Number of trees = 100
Number of trees = 500
Number of trees = 1000

- A) 1~2~3
- B) 1<2<3
- C) 1>2>3
- D) None of these

When you use the boosting algorithm you always consider the weak learners. Which of the following is the main reason for having weak learners? 1 point
To prevent overfitting
To prevent under fitting

- A) 1
- B) 2
- C) 1 and 2
- D) None of these

Which of the following is a widely used and effective machine learning algorithm based on the idea of bagging?

1 point

- Decision Tree
- Regression
- Random Forest
- Classification

The most widely used metrics and tools to assess a classification model are:

1 point

- Confusion matrix
- Cost-sensitive accuracy
- Area under the ROC curve
- All of the above

Which of the following is a good test dataset characteristic?

1 point

- Large enough to yield meaningful results
- Is representative of the dataset as a whole
- Both A and B
- None of the above

What is `pca.components_` in Sklearn?

2 points

- Set of all eigen vectors for the projection space
- Matrix of principal components
- Result of the multiplication matrix
- None of the above options

How can you prevent a clustering algorithm from getting stuck in bad local optima?

1 point

- Set the same seed value for each run
- Use multiple random initializations
- Both A and B
- None of the above

Which of the following is a reasonable way to select the number of principal components "k"?

1 point

- Choose k to be the smallest value so that at least 99% of the variance is retained.
- Choose k to be 99% of m ($k = 0.99 * m$, rounded to the nearest integer).
- Choose k to be the largest value so that 99% of the variance is retained.
- Use the elbow method

Suppose you have trained a logistic regression classifier and it outputs a new example x with a prediction $h_0(x) = 0.2$. This means

2 points

- Our estimate for $P(y=0 | x)$
- Our estimate for $P(y=1 | x)$
- Our estimate for $P(y=0 | x)$
- All the above

Which of these problems does not fall into 3 main types of ML tasks: classification, regression, and clustering?

1 point

- Identifying a topic of a live-chat with a customer [+]
- Grouping news into topics
- Predicting LTV (Life-Time Value) - the amount of money spent by a customer in a certain large period of time
- Listing top products that a user is prone to buy (based on his/her click history)

This content is neither created nor endorsed by Google.

Google Forms