

CS564 Mid-Term MCQ

maheeth2000@gmail.com [Switch accounts](#)

 Draft saved

***Required**

Email *

maheeth2000@gmail.com

Please Enter your name *

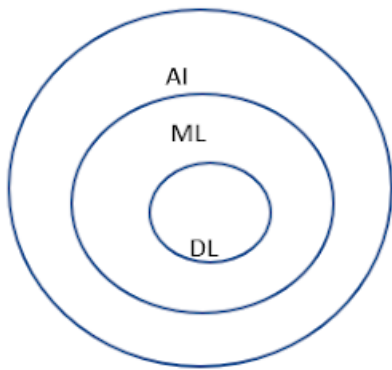
Maramreddy Maheeth Reddy

Please Enter your roll *

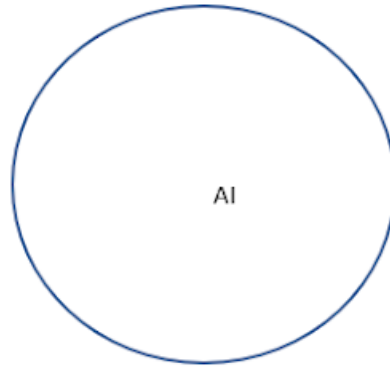
1801CS31



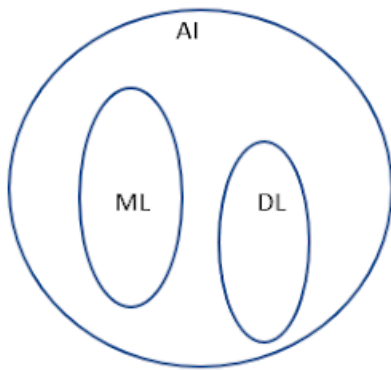
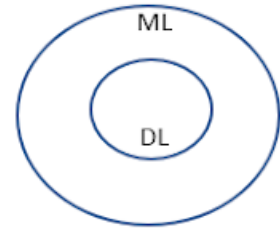
1. If AI stands for Artificial Intelligence, ML stands for Machine Learning and DL stands for Deep Learning, then which figure correctly represents the relation between AI, ML and DL.



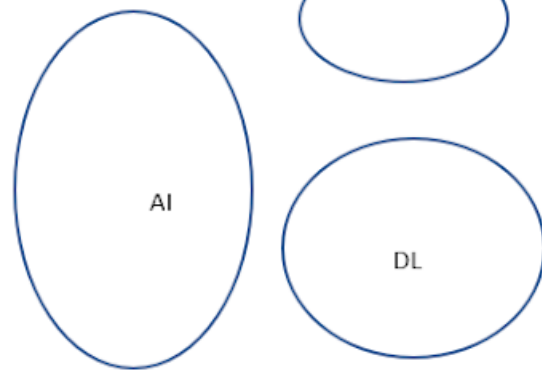
(A)



(B)



(C)



(D)

☒ (A)

☐ (B)

☐ (C)

☐ (D)

Clear selection



2. Which of the following does not relate to numerical functions in the various function representation of machine learning algorithms?

- ☒ Case-based
- ☐ Neural network
- ☐ Linear regression
- ☐ None

Clear selection

3. Machine Learning solves which of the following problems:

- ☐ Identification of patient's clusters from breast cancer dataset based on the genetic details of the patients.
- ☐ Prediction of the profit margin of the company over the period of time.
- ☐ Only first option
- ☒ Both first and second options

Clear selection

4. Find the odd one out.

- ☒ PCA
- ☐ LDA
- ☐ Naïve Bayesian
- ☐ Linear regression

Clear selection



5. Which of the following is an example of clustering?

- ☐ Structuring search results
- ☐ Suggesting related pages
- ☐ Automatic directory construction/update
- ☒ All of the above
- ☐ None of the above

Clear selection

6. Which of the following can act as possible termination conditions in K-Means?

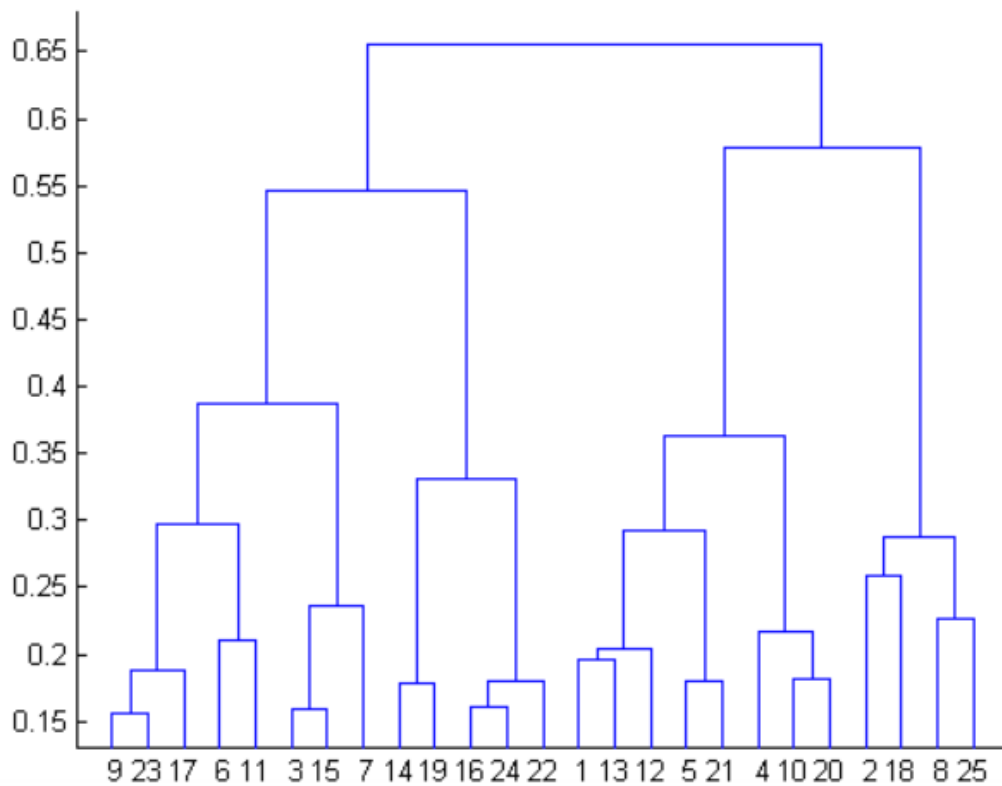
(i) For a fixed number of iterations. (ii) Assignment of observations to clusters does not change between iterations. Except for cases with a bad local minimum. (iii) Centroids do not change between successive iterations. (iv) Terminate when RSS falls below a threshold.

- ☐ i, iii and iv
- ☐ i, ii and iii
- ☐ i, ii and iv
- ☒ i, ii, iii and iv

Clear selection



7. You discovered the following dendrogram after running K-Means Clustering on a dataset. Which of the following can be inferred from the dendrogram's findings?



- ☐ There were 28 data points in clustering analysis.
- ☐ The best no. of clusters for the analysed data points is 4.
- ☐ The proximity function used is Average-link clustering.
- ☒ The above dendrogram interpretation is not possible for K-Means clustering analysis.

Clear selection



8. Change in either of Proximity function, no. of data points or no. of variables while using agglomerative clustering algorithm for the same dataset will lead to:

- ☒ Different clustering results and different dendrograms
- ☐ Different clustering results and same dendrograms
- ☐ Same clustering results and different dendrograms.
- ☐ Same clustering results and same dendrograms

Clear selection

9. The K-means clustering fails to give good results. What could be the possible reasons? (i) Data points with outliers (ii) Data points with different densities (iii) Data points with round shapes (iv) Data points with non-convex shapes

- ☐ i and ii
- ☐ ii and iii
- ☐ ii and iv
- ☒ i, ii and iv
- ☐ i, ii, iii and iv

Clear selection

10. What is the advantage of the k-Medoids Clustering Algorithm over the k-Means Clustering (Lloyd's) Algorithm?

- ☐ Uses iterative refinement
- ☒ More resistant to outliers
- ☐ Represents clusters by centers
- ☐ All of the above

Clear selection

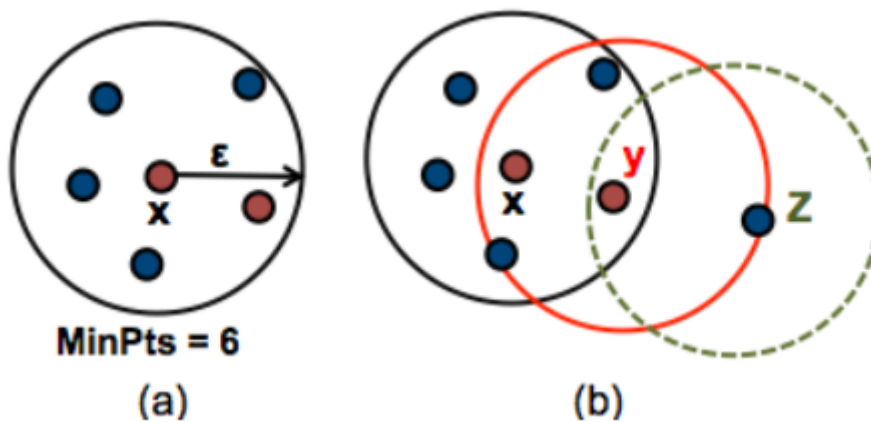


11. The K-medoid clustering fails to give good results. What could be the possible reasons? (i) Data points with outliers (ii) Data points with different densities (iii) Data points with round shapes (iv) Data points with non-convex shapes

- ☐ i and ii
- ☐ ii and iii
- ☒ ii and iv
- ☐ i, ii and iv
- ☐ i, ii, iii and iv

Clear selection

12. After applying the DBSCAN Algorithm on a dataset, we get the following clusters (as shown below in the figure). Identify the core point, border point, and noise point.



- ☒ Core point - x , Border point - y, Noise point - z
- ☐ Core point - x, Border point - y, No Noise point
- ☐ No Core point, Border point - x and y, Noise point - z
- ☐ Core point - x and y , Border point - z, No Noise point

Clear selection

13. The worst-case time complexity of single-link, complete-link and average-link hierarchical clustering are:

- ☐ $O(n^2 \log n)$, $O(n^2)$, $O(n^2 \log n)$
- ☒ $O(n^2 \log n)$, $O(n^2 \log n)$, $O(n^2 \log n)$
- ☐ $O(n^2)$, $O(n^2)$, $O(n^2 \log n)$
- ☐ $O(n^2)$, $O(n^2)$, $O(n^2)$

Clear selection

14. Which of the following is not the property of Jaccard coefficient?

- ☐ Measures the fraction of true positives while ignores the true negatives
- ☐ Jaccard coefficient of a perfect cluster indicates that there are no false positives nor false negatives
- ☒ It's a measure of dis-similarity for the two sets of data.
- ☐ The Jaccard coefficient is a measure of the percentage of overlap between sets.

Clear selection

15. Which of the following is correct about cohesion and separation?

- ☐ Cohesion measures the closeness between two clusters.
- ☐ Separation measures the separation between two clusters.
- ☒ Both
- ☐ None

Clear selection



16. Students in IIT Patna were asked if they have ever driven after drinking. They also were asked, "How many days per month do you drink at least two beers?" In the following discussion, P = the probability a student says "yes" they have driven after drinking. This is modelled using X = days per month of drinking two beers and the coefficient α and β for the modelled function $(\alpha + \beta X)$ are -1.5514 and 0.19031 , respectively. If a student drinks 4 days per month, then find the probability of ever having driven after drinking.

- ☐ 79.9%
- ☐ 90.2%
- ☒ 31.2%
- ☐ 54.5%

Clear selection

17. Which of the following statement is incorrect?

- ☐ The method for calculating loss function in linear regression is the mean squared error whereas for logistic regression it is maximum likelihood estimation.
- ☐ Linear Regression is used to handle regression problems whereas Logistic regression is used to handle the classification problems.
- ☐ Linear regression provides a continuous output but Logistic regression provides discrete output.
- ☒ Logistic regression uses Sigmoid function but not logistic function.

Clear selection



18. Which of the following is/are valid iterative strategies for treating missing values before clustering analysis?

- ☐ Imputation with mean
- ☐ Nearest Neighbor assignment
- ☒ Imputation with Expectation Maximization algorithm
- ☐ All of the above

Clear selection

19. The most popularly used dimensionality reduction algorithm is Principal Component Analysis (PCA). Which of the following is/are true about PCA? 1) PCA is an unsupervised method. 2) It searches for the directions that data have the largest variance. 3) Maximum number of principal components \leq number of features. 4) All principal components are orthogonal to each other

- ☐ 1 and 2
- ☐ 2 and 3
- ☐ 1 and 4
- ☒ All of the above

Clear selection

20. In which of the following case LDA will fail?

- ☒ If the discriminatory information is not in the mean but in the variance of the data
- ☐ If the discriminatory information is in the mean but not in the variance of the data
- ☐ If the discriminatory information is in the mean and variance of the data
- ☐ None of these

Clear selection



21. In PCA, what will happen when eigenvalues are roughly equal?

- ☐ PCA will perform outstandingly
- ☒ PCA will perform badly
- ☐ Can't say
- ☐ None of the above

Clear selection

22. How do we perform Bayesian classification when some features are missing?

- ☐ By assuming the missing values as the mean of all values.
- ☐ We ignore the missing features.
- ☒ We integrate the posteriors probabilities over the missing features.
- ☐ Drop the features completely.

Clear selection

23. Which of the following can be the first 2 principal components after applying PCA? 1) (0.5, 0.5, 0.5, 0.5) and (0.71, 0.71, 0, 0) 2) (0.5, 0.5, 0.5, 0.5) and (0, 0, -0.71, -0.71) 3) (0.5, 0.5, 0.5, 0.5) and (0.5, 0.5, -0.5, -0.5) 4) (0.5, 0.5, 0.5, 0.5) and (-0.5, -0.5, 0.5, 0.5)

- ☐ 1 and 2
- ☐ 1 and 3
- ☐ 2 and 3
- ☒ 3 and 4

Clear selection



24. From a standard deck of playing cards, a single card is drawn. The probability that the card is king is $\frac{4}{52}$, then calculate posterior probability $P(\text{King}|\text{Face})$, which means the drawn face card is a king card.

- ☐ $\frac{2}{3}$
- ☒ $\frac{1}{3}$
- ☐ $\frac{1}{5}$
- ☐ $\frac{1}{8}$

Clear selection

25. A doctor is aware that disease meningitis causes a patient to have a stiff neck, and it occurs 60% of the time. He is also aware of some more facts, which are given as follows: The Known probability that a patient has meningitis disease is $\frac{1}{50,000}$. The Known probability that a patient has a stiff neck is 5%.

- ☒ 0.00024
- ☐ 0.01133
- ☐ 0.00005
- ☐ 0.03333

Clear selection

26. The previous probabilities in Bayes Theorem that are changed with the help of new available information are classified as

- ☐ independent probabilities
- ☒ posterior probabilities
- ☐ interior probabilities
- ☐ dependent probabilities

Clear selection



27. The method in which the previously calculated probabilities are revised with new probabilities is classified as

- ☐ Updating theorem
- ☐ Revised theorem
- ☒ Bayes theorem
- ☐ Dependency theorem

Clear selection

28. A feature F1 can take certain value: A, B, C, D, E, & F and represents grade of students from a college. Which of the following statement is true in following case?

- ☐ Feature F1 is an example of nominal variable.
- ☒ Feature F1 is an example of ordinal variable.
- ☐ It doesn't belong to any of the above category.
- ☐ Both of these

Clear selection

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

- ☒ Choose k to be the smallest value so that at least 99% of variance is retained.
- ☐ Choose k to be 99% of m.
- ☐ Choose k to be the largest value so that at least 99% of variance is retained.
- ☐ Use elbow method

Clear selection



30. Which of the following is an example of feature extraction?

- ☐ Constructing bag of words vector from an email.
- ☐ Applying PCA projects to a large high-dimensional data
- ☐ Removing stopwords in a sentence
- ☒ All of the above

Clear selection

31. In each schema, the set of bit strings containing the indicated as

- ☐ 0s and 1s
- ☐ only 0s
- ☐ only 1s
- ☒ 0s, 1s, *s

Clear selection

32. Disadvantages of Naive Bayes Classifier are

- ☒ Naive Bayes assumes that all features are independent or unrelated, so it cannot learn the relationship between
- ☐ It performs well in Multi-class predictions as compared to the other
- ☐ Naïve Bayes is one of the fast and easy ML algorithms to predict a class
- ☐ It is the most popular choice for text classification problems.

Clear selection



33. 0^*10 represents the set of bit strings that includes exactly

- ☒ 0010, 0110
- ☐ 0010, 0010
- ☐ 0100, 0110
- ☐ 0100, 0010

Clear selection

34. Which of the following is true about stopping of Genetic Algorithm? 1. Stopping algo after k_1 generations 2. Stopping algo post the occurrence of at least k_2 generations, at least for k_1 iterations the max(or avg) value of J for all population members has risen by ϵ and not more 3. Stopping algo the moment J takes on a value higher than a fixed value

- ☐ 1
- ☐ 2
- ☐ 3
- ☒ all 3

Clear selection

35. Which of the following are features of optimisation problems? a) a set of parameters capable of adjustments b) an objective function that pools the parameters into a one single value c) a set of constraints on the parameters

- ☐ a and b
- ☐ a and c
- ☒ a,b,c
- ☐ b and c

Clear selection



36. Which of the following is true about EM algorithm? 1. Given current parameters $\theta(t)$ and observed data, obtain $Q(\theta|\theta(t))$ 2. Maximise expectation $Q(\theta|\theta(t))$ as $\theta(t+1) = \operatorname{argmax} Q(\theta|\theta(t))$

- ☐ 1
- ☐ 2
- ☐ None
- ☒ Both

Clear selection

37. If optimisation problem is to maximise $J(\theta) \geq 0$, then which is the correct way to express the cost function? I. $J(\theta) = 1/J_bar(\theta) + \epsilon$ II. $J(\theta) = -J_bar(\theta) + \max |J_bar(\theta)|$ where ϵ is a small +ve number.

- ☒ Only I
- ☐ Only II
- ☐ Both I and II
- ☐ None of these

Clear selection



38. Considering GA, which of these are valid sequences? 1. Encoding -> Choosing Fitness Function -> Initialise population -> Selection -> Crossover -> Mutation 2. Initialise population -> Encoding -> Choosing Fitness Function -> Selection -> Crossover -> Mutation 3. Encoding -> Initialise population -> Choosing Fitness Function -> Selection -> Crossover -> Mutation 4. Choosing Fitness Function -> Encoding -> Initialise population -> Selection -> Crossover -> Mutation

- ☐ 1 and 4
- ☐ 1 and 2
- ☒ 2 and 3
- ☐ 3 and 4

Clear selection

39. What should be minimum possible value of minPoints in DBSCAN?

- ☒ minPoints \geq Dimensions of the dataset + 1.
- ☐ minPoints \geq Dimensions of the dataset.
- ☐ Depends upon the data distribution.
- ☐ minPoints > 0 .

Clear selection



40. In DBSCAN, which is not correct about epsilon?

- ☐ The value of epsilon can be decided from the K-distance graph.
- ☐ The point of maximum curvature (elbow) in this graph tells us about the value of epsilon
- ☐ If the value of epsilon chosen is too small then a higher number of clusters will be created
- ☒ If the value of epsilon chosen is too large then more data points will be taken as noise

Clear selection

41.

Find the value of x such that fitness function $f(x) = 15x - x^2$ where $x \geq 0$ and $x \leq 15$ is maximised.

- ☒ 7.5
- ☐ 15
- ☐ 0
- ☐ 1

Clear selection

Page 1 of 1

Submit

Clear form

Never submit passwords through Google Forms.

This content is neither created nor endorsed by Google. [Report Abuse](#) - [Terms of Service](#) - [Privacy Policy](#).

Google Forms

