Unit Test 2- Machine Learning

Marks: 30   Duration: 1 Hour

$\cdots$

Points: **28/30**

1. RollNo *

> 41310

2. Name *

> Prem Vinod Bansod

3. ----------Clustering starts from the bottom and proceeds by merging the clusters until a stop criterion is reached.(Unit 6, CO4(H),Bloom's level(Understand)) *
(1/1 Point)

- ○ d. k-means
- ○ a. Divisive
- ◉ b. Agglomerative ✓
- ○ c. Both a and b

4. Which classification algorithm is used when there is a case of a feature can be present or absent. (binary distribution) (Unit 4, CO3(M),Bloom's level(Understand)) *
(1/1 Point)

- ⦿ Bernoulli naive bayes ✓

- ◯ multinomial naive bayes

- ◯ Gaussian naive bayes

- ◯ None

5. What is the type of Hierarchical Clustering (Unit 5, CO4(L),Bloom's level(Remember)) *
(1/1 Point)

- ◯ k-mediod

- ◯ K-means

- ◯ Dendogram

- ⦿ Top-Down Clustering (Divisive) ✓

6. Which of these are reasons for Deep Learning recently taking off? (Unit 6, CO5(H),Bloom's level(Understand)) *
(1/1 Point)

- ☑ We have access to a lot more data. ✓

- ☑ We have access to a lot more computational power ✓

- ☐ Neural Networks are a brand new field.

- ☑ Deep learning has resulted in significant improvements in important applications such as online advertising, speech recognition, and image recognition. ✓

7. In which of the following scenario a gain ratio is preferred over Information Gain?
   (Unit 5, CO3(M),Bloom's level(Analyze)) *
   (1/1 Point)

   ◉ When a categorical variable has very large number of category ✓

   ○ None of these

   ○ When a categorical variable has very small number of category

   ○ Number of categories is the not the reason

8. The time complexity of hierarchical clustering is (Unit 4, CO4(H),Bloom's
   level(analysis)) *
   (1/1 Point)

   ○ o(n)

   ◉ O(n^2) ✓

   ○ n logn

   ○ logn

9. Threshold for binarizing (mapping to booleans) of sample features in BernalliNB
   function in Skikit learn is set using (Unit 4, CO3(H),Bloom's
   level(Implement/design)) *
   (1/1 Point)

   ◉ binarize ✓

   ○ None

   ○ class_prior

   ○ alpha

10. Decision trees that are too large are susceptible to a phenomenon known as (Unit 5, CO3(H),Bloom's level(Understand)) *
(-/1 Point)

○ A) Overfitting

○ B) Pruning

○ None of the above

○ c) Both A and B

11. which type of rating can be collected constantly and do not require additional efforts from the side of the user (Unit 6, CO5(L),Bloom's level(Understand) *
(1/1 Point)

◉ implicit ratings ✓

○ explicit ratings

12. Match the pair (Unit 4, CO3(M),Bloom's level(Understand))

a. SVM   1.Nondeterministic algorithm
b. NN   2. Deterministic algorithm *
(1/1 Point)

◉ a-2, b-1 ✓

○ a-1, b-2

13. In the Random forest algorithm, The final prediction is calculated by_____ the predictions from all decision trees. (Unit 5, CO3(H),Bloom's level(design)) *
(1/1 Point)

○ Taking Median of

○ taking Mode Of

○ adding

◉ averaging ✓

14. In normal distribution (Unit 4, CO3(H),Bloom's level(Analyze)) *
    (1/1 Point)

    ○ Mode>Mean>Median

    ○ Median>Mean>Mode

    ◉ Mode=Median=Mean ✓

    ○ None

15. The raw data are first processed offline like item based
    filtering or some dimensionality reduction techniques. At run time, only
    the precomputed or "learned" model is required to make predictions in
    recommendation system. This is called: (Unit 6, CO5(M),Bloom's
    level(understand)) *
    (1/1 Point)

    ○ pre-processing based

    ○ memory-based

    ○ none

    ◉ model-based ✓

16. Which classification algorithm is used when there is a case of a continuous
    distribution characterized by its mean and variance.(Unit 4, CO3(M),Bloom's
    level(Understand)) *
    (1/1 Point)

    ◉ Gaussian naive bayes ✓

    ○ None

    ○ multinomial naive bayes

○ Bernoulli naive bayes 2

17. The difference of SVR and SVM is: (Unit 4, CO3(M),Bloom's level(Understand)) *
(1/1 Point)

○ margin of tolerance is not used

○ real no

◉ margin of tolerance is used ✓

○ Both same

18. Fill in the blanks

the Normal curve is _____ and the total AUC is _____. (Unit 4,
CO3(M),Bloom's level(Analyze)) *
(1/1 Point)

○ symmetrical,<1

○ Nonsymmetrical, >1

○ Nonsymmetrical, 1

◉ symmetrical,1 ✓

19. Item1 Item2 Item3 Item4 Item5
Alice 1 3 3 2 ?
User1 2 4 2 2 4
User2 1 3 3 5 1
User3 4 5 2 3 3
User4 1 1 5 2 1

Predict the rating of Item 5 to Use Alice by naive bayes classifier (Unit 6,
CO5(H),Bloom's level(Solve)) *
(1/1 Point)

◉ 1 ✓

○ 2

○ 3

○ None of the above

20. Naive Bayes classifier assumes that the presence of a particular feature in a class is related to the presence of any other feature. (Unit 6, CO4(H),Bloom's level(Understand)) *
(1/1 Point)

⦿ False ✓

○ True

21. …….. syntax to import agglomerative clustering.(Unit 6, CO4(H),Bloom's level(design)) *
(-/1 Point)

⦿ a. sklearn.cluster import AgglomerativeClustering

○ b. sklearn import AgglomerativeClustering

○ c. sklearn.hierarchical import AgglomerativeClustering

○ d. scipy.cluster import AgglomerativeClustering

22. Which of the following would have a constant input in each epoch of training a Deep Learning model?
(Unit 6, CO5(H),Bloom's level(Understand)) *
(1/1 Point)

⦿ Weight between input and hidden layer ✓

○ Weight between hidden and output layer

○ Biases of all hidden layer neurons

○ Activation function of output layer

23. ------ is a type of Bagging Algorithm. (Unit 5, CO3(M),Bloom's level(remember)) *
    (1/1 Point)

    ○ GBM

    ○ Light GBM

    ○ XGBM

    ◉ Random forest ✓

24. In SVM the complexity of trained classifier is characterized by (Unit 4,
    CO3(L),Bloom's level(Understand/remember)) *
    (1/1 Point)

    ○ None of the above

    ◉ the # of support vectors ✓

    ○ the dimensionality of the data

25. Which of the following algorithm doesn't uses learning Rate as of one of its
    hyperparameter?
    (Unit 5, CO3(M),Bloom's level(Understand)) *
    (1/1 Point)

    ○ Gradient Boost

    ○ None of the above

    ○ AdaBoost

    ◉ Random Forest ✓

26.

	Item1	Item2	Item3	Item4	Item5
User1	3	1	2	3	3
User2	4	3	4	3	5
User3	3	3	1	5	4
User4	1	5	5	2	1

Compute Cosine similarity between item 1 and item 5. (Unit 6, CO5(M),Bloom's level(Solve)) *
(1/1 Point)

○ None

○ 1

◉ 0.99 ✓

○ 0.92

27. Which of the following is/are not true about DBSCAN clustering algorithm: 1. For data points to be in a cluster, they must be in a distance threshold to a core point 2. It has strong assumptions for the distribution of data points in dataspace 3. It has substantially high time complexity of order O(n3) 4. It does not require prior knowledge of the no. of desired clusters 5. It is robust to outliers

(Unit 5, CO4(H),Bloom's level(Understand)) *
(1/1 Point)

◉ 2 and 3 ✓

○ 2 only

○ 4 only

○ 1 only

28. MultinomialNB(alpha=1.0, class_prior=None, fit_prior=True)
Here what is alpha indicate (Unit 4, CO3(H),Bloom's level(Implement/design)) *
(1/1 Point)

◉ smoothing parameter ✓

○ None

○ Threshold

○ Weight

29. Male height: Normal with μ = 70.0″ and σ = 2.8″ Find the height range of 68%
(Unit 4, CO3(H),Bloom's level(Analyze/solve)) *
(1/1 Point)

○ 67.2 to 70

◉ 67.2 to 72.8 ✓

○ 60 to 70

○ 65.2 to 70.8

30. Point out the correct statement.
(Unit 5, CO4(H),Bloom's level(Remember)) *
(1/1 Point)

○ In general, the merges and splits are determined in a greedy manner

○ The choice of an appropriate metric will influence the shape of the clusters

◉ All of the mentioned ✓

○ Hierarchical clustering is also called HCA

31. Which of the following is a method of choosing the optimal number of clusters for k-means?
(Unit 5, CO4(M),Bloom's level(Remember)) *
(1/1 Point)

◉ all of the above ✓

○ the elbow method

○ the silhouette method

○ cross-validation

32. Which of the following true about weak learners used in the ensemble model?

    1. They have low variance and they don't usually overfit
    2. They have high bias, so they can not solve hard learning problems
    3. They have high variance and they don't usually overfit
    Unit 5, CO3(H),Bloom's level(understand)) *
    (1/1 Point)

◉ 1 and 2 ✓

○ 2 and 3

○ None of these

○ 1 and 3

Go back to thank you page

Powered by Microsoft Forms | Privacy and cookies | Terms of use