Bishop Heber College (Autonomous)- Tiruchirappalli - 17

END SEMESTER EXAMINATIONS - APRIL 2022

M.Sc. Data Science **Course Code: P21DS206**

PRACTICAL MACHINE LEARNING

Duration: 3 Hrs. **Max. Marks : 75**

Section – A **(20 x1= 20)**

Answer all the questions:

Choose the correct answer:

1. Which of the following is a classification application?
   a. Sales prediction b. Character recognition c. Temperature forecast
   d. Population increase
2. Identify the kind of learning algorithm for "written alphabet identification".
   a. Prediction b. Pattern recognition c. Association analysis d. Clustering
3. In a neural network, the amount of output of one unit received by another unit depends on
   a. number of layers b. output units c. activation value d. weights
4. The main difference between Adaline and Perceptron
   a. Class labels in Perceptron and continuous values in Adaline
   b. Class labels in Adaline and continuous values in Perceptron
   c. Linear decision boundary in Perceptron and non linear boundary in Adaline
   d. Threshold function is used in Adaline while it is not used in Perceptron
5. Suppose you have identified two multi-collinear features. Which of the following is NOT a solution?
   a. Remove the two multi-collinear features b. Remove any one feature
   c. Use Ridge regression d. Use Lasso Regression
6. If two variables "study hours" and "marks" are positively correlated in a dataset, what can we say about their covariance value?
   a. Covariance > 0 b. Covariance < 0 c. Covariance = 0 d. Covariance <= 0
7. Given an estimation regression equation with one predictor with coefficient as 4.67 and intercept as 2.38. What is the predicted y value when predictor value is 1.85?
   a. 11.02 b. 12.96 c. 9.073 d. 10.15
8. Which of the following is NOT an attribute selection measure in Decision tree?
   a. Information gain b. Gini index c. Gain ratio d. Majority vote
9. What does PCA perform?
   a. Missing data processing b. Noisy data detection
   c. Feature subset selection d. Reduce dimensions
10. Why is so important standardize data in PCA?
    a. Faster training time b. Better interpretation of data
    c. Find the features with high variance d. Data wrangling

11. A data set uses Movie name, Genre of movie, budget amount and the language as attributes. Genre of movie shall be termed as

a. Numeric b. Categorical c. Ordinal d. Interval scaled

12. In a data set with values <1000,2000,3000,9000>, what is the new value of 3000 after min max scaling?

a. 0.125 b. 0.25 c. 05 d. 0.75

13. Overfitting happens due to -

a. High bias b. Complex model c. Less variance d. Larger data set

14. Which of the following is true about bagging among the three given statements?

S1. Bagging can be done in parallel S2. Reduces bias S3. addresses overfitting .

a. S1 and S2 b. S2 and S3 c. S1 and S3 d. S1, S2 and S3

15. What is the formula for calculating accuracy of a classification model?

a. (TP+ TN)/total no. of records b. TP/(TP + FP) c. TP/(TP+FN) d. TN/(TP+FN)

16. If you ensemble 3 models with 70% accuracy with majority voting method. What is the minimum accuracy obtained?

a. <70% b. =70% c. > 70% d. 100%

17. What is the range of sigmoid activation function in a neural network ?

a. -1 to 1 b. 0 to 1 c. – infinity to infinity d. 0 to infinity

18. Which of the following clustering algorithm is most sensitive to outliers?

a. k-means b. k-medians c. k-modes d. k-medoids

19. Which of the following are true?

1. Clustering analysis is negatively affected by multicollinearity of features
2. Clustering analysis is negatively affected by heteroscedasticity

a. 1 only b. 2 only c. 1 and 2 d. Neither is true

20. Give the Manhattan distance between two points (3,5) and (7,6)

a. 5 b. 17 c. 4.123 d. (4,1)

SECTION - B **5 x 5 = 25**

Answer ALL questions

21.a. Discuss the types of Machine Learning and the steps involved in Machine learning process.

(or)

21.b. Explain the basic structure of Adaline neural network with a simple example.

22. a. Demonstrate the different kernel used in SVM and explain how the kernel is used in classifying non separable data.

(or)

22.b. Given a logit function z= 0.001 + 0.6*x1-0.74*x2 and given four data points for the classification (Yes/No). Classify the four data points as Yes or No

No.	X1	X2
1	20	10
2	4	5
3	7	6
4	7	2

23.a. Explain the methods to treat missing data and standardizing in a numeric data set.

(or)

23.b.Present the steps to be followed in Linear Discriminant Analysis of a data set.

24.a.Explain the classification model evaluation with confusion matrix

(or)

24.b.Demonstrate the technique of Ada Boosting with an example

25.a. Illustrate the process in Density based hierarchical clustering.

(or)

25.b.Demonstrate the technique of k-means clustering with an example.

SECTION C **3 x 10 = 30**

Answer ANY THREE

26.Write a detailed report on applying Perceptron neural network to a classification problem.

27. Given a data set with 10 records. Apply the information gain procedure to decide the first two levels of the decision tree. (Classification Label : Liked)

Student	Prior Experience	Course	Time	Liked
1	Yes	Programming	Day	Yes
2	No	Programming	Day	No
3	Yes	History	Night	No
4	No	Programming	Night	Yes
5	Yes	English	Day	Yes
6	No	Programming	Day	No
7	Yes	Programming	Day	No
8	Yes	Mathematics	Night	Yes
9	Yes	Programming	Night	Yes
10	Yes	Programming	Night	No

28. Given the data for five students: Use PCA method to reduce the dimension.

Student	Math	English	Art
1	90	60	90
2	90	90	30
3	60	60	60
4	60	60	90
5	30	30	30

29. Illustrate any two ensemble methods of improving the predictive or classification model

30. Demonstrate the organization of feed forward network and how learning happens in a feed forward network.

****