"Techno-Social Excellence"
Marathwada Mitramandal's
INSTITUTE OF TECHNOLOGY (MMIT)
Lohgaon, Pune-411 047

"Towards ubiquitous Computing Technology"

Department of Computer Engineering

End Sem Exam Solution

Programme: BE **Semester: II**

Course Code: 410250 **Course Name: ML**

Branch: Computer Engineering **Academic Year: 2019-20**

Duration: 1 Hour **Max Marks: 50**

Student PRN No.

Student Roll No.

Instructions:

Questions are of MCQs type. Question paper will be as per the format
a) Maximum Marks- 50 questions
b) Total question –30
c) 1 Mark MCQ-16question
d) 2 Marks MCQ/short questions – 11 questions
e) 4 Marks MCQ/short answer questions – 03questions.
f) For numerical based subject 2MarksMCQis also based on short numerical.

1.Previous probabilities in Bayes Theorem that are changed with help of new available information are classified as _____ [1 M]
a) independent probabilities
b) posterior probabilities
c) interior probabilities
d) dependent probabilities

2. The method in which the previously calculated probabilities are revised with new probabilities is classified as [1 M]
a] updating theorem
b] revised theorem
c] Bayes theorem
d]dependency theorem

3. The previous probabilities in Bayes Theorem that are changed with the help of new available information are classified as [1 M]
a) independent probabilities

b) posterior probabilities
c) interior probabilities
d) dependent probabilities

4. The model which assumes that all our features are binary such that they take only two values is [1 M]
a) Multinomial Naïve Bayes
b) Gaussian Naïve Bayes
c) Bernoulli Naïve Bayes
d) none

5. The effectiveness of an SVM depends upon: [1 M]
a) Selection of Kernel
b) Kernel Parameters
c) Soft Margin Parameter C
d) All of the above

6. In Classification Model, Which Technique can help you to choose a threshold that balance sensitivity and specificity [1 M]
a) Confusion Matrix
b) ROC curve
c) MAPE
d) None of the Above

7. In Decision Tree, by comparing the impurity across all possible splits in all possible Predictors, the next split is choosen. How we can measure the Impurity ? [1 M]
a) UC
b) Entropy, Ginni-Index
c) ROC
d) MAPE

8. How we can avoid the overfitting in Decision Tree [1 M]
c) Both of above
a) CHAID(Stopping the Tree Growth
b) Pruning the Full Grown Tree
d) None of the Above

9. Predictive Errors are due to [1 M]
a) Bias
b) Variance
c) Both of above
d) None of the Above

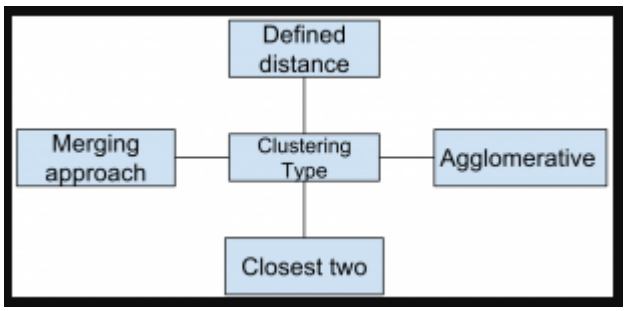
10. Any linear model can be turned into a non-linear model by applying the kernel trick to the model. [1 M]

a) true
b) false

11. Random Forest Modeling (Ensemble Modeling) uses . [1 M]

a) Bagging(BootStrap Samples)
b) Boosting
c) Both of above
d) None of the Above

12. Which of the following clustering type has characteristic shown in the below Figure? [1 M]



a) Partitional
b) Hierarchical
c) Naive bayes
d) None of the mentioned
Explanation: Hierarchical clustering groups data over a variety of scales by creating a cluster tree or dendrogram.

13. Point out the correct statement. [1 M]
a) The choice of an appropriate metric will influence the shape of the clusters
b) Hierarchical clustering is also called HCA
c) In general, the merges and splits are determined in a greedy manner
d) All of the mentioned
Explanation: Some elements may be close to one another according to one distance and farther away according to another.

14. Which of the following is finally produced by Hierarchical Clustering? [1 M]
a) final estimate of cluster centroids
b) tree showing how close things are to each other
c) assignment of each point to clusters

d) all of the mentioned**.**
Explanation: Hierarchical clustering is an agglomerative approach.

15. Which of the following is required by K-means clustering? [1 M]
a) defined distance metric
b) number of clusters
c) initial guess as to cluster centroids
d) all of the mentioned
Explanation: K-means clustering follows partitioning approach.

16. Point out the wrong statement. [1 M]
a) k-means clustering is a method of vector quantization
b) k-means clustering aims to partition n observations into k clusters
c) k-nearest neighbor is same as k-means
d) none of the mentioned
Explanation: k-nearest neighbor has nothing to do with k-means.

17. Which of the following clustering requires merging approach? [2 M]
a) Partitional
b) Hierarchical
c) Naive Bayes
d) None of the mentioned
Explanation: Hierarchical clustering requires a defined distance as well.

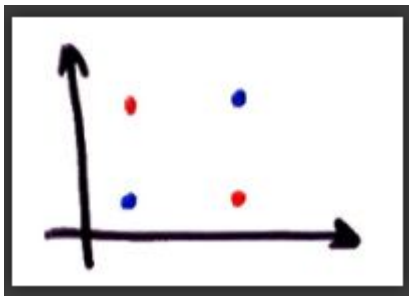
18. a system that is capable of predicting the future preference of a set of items for a user, and recommend the top items. [2 M]
a) Recommendation Systems
b) collaborative filtering
C) Content based Systems
d)all of the above

19. A content based recommender works with data that the user provides, either explicitly (rating) or implicitly. [2 M]

a)true

b)false

20. **Is the data linearly separable?** [2M]



a) Yes
b) No
Explanation :If you can draw a line or plane between the data points, it is said to be linearly separable.

21. Which of the following are universal approximators? [2M]

a) Kernel SVM
b) Neural Networks
c) Boosted Decision Trees
d) All of the above

Explanation :All of the above methods can approximate any function.

22. Decision Tree is a display of an algorithm. [2M]

a) True
b) False

23. A _____ is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. [2M]
a) Decision tree
b) Graphs
c) Trees
d) Neural Networks

24. Choose from the following that are Decision Tree nodes? [2M]
a) Decision Nodes
b) End Nodes
c) Chance Nodes
d) All of the mentioned

25. . Which of the following is/are true about bagging trees? [2M]

1. In bagging trees, individual trees are independent of each other

2. Bagging is the method for improving the performance by aggregating the results of weak learners

a) 1
b) 2
c) 1 and 2
d) None of these

Explanation:Both options are true. In Bagging, each individual trees are independent of each other because they consider different subset of features and samples.

26 . Which of the following is/are true about Random Forest and Gradient Boosting ensemble methods? [2M]

1. Both methods can be used for classification task
2. Random Forest is use for classification whereas Gradient Boosting is use for regression task
3. Random Forest is use for regression whereas Gradient Boosting is use for Classification task
4. Both methods can be used for regression task

a) 1
b) 2
c) 3
d) 1&4

Explanation: Both algorithms are design for classification as well as regression task

27. Which of the following algorithm doesn't uses learning Rate as of one of its hyperparameter? [2M]

1. Gradient Boosting
2. Extra Trees
3. AdaBoost
4. Random Forest

a) 1 and 3
b) 1 and 4
c) 2 and 3
d) 2 and 4

Explanation: Random Forest and Extra Trees don't have learning rate as a hyperparameter

28. Which of the following algorithm are not an example of ensemble learning algorithm? [4 M]

a) Random Forest
b) Adaboost
c) Extra Trees
d) Decision Trees

Explanation: Decision trees doesn't aggregate the results of multiple trees so it is not an ensemble algorithm

29. Which of the following splitting point on feature x1 will classify the data correctly? [4]

a) Greater than x11
b) Less than x11
c) Equal to x11
d) None of above

Explanation: If you search any point on X1 you won't find any point that gives 100% accuracy

30. What will be the minimum accuracy you can get? [4]

a) Always greater than 70%
b) Always greater than and equal to 70%
c) It can be less than 70%
d) None of these

Explanation: Refer below table for models M1, M2 and M3.

Actual predictions	M1	M2	M3	Output
1	1	0	0	0
1	1	1	1	1
1	1	0	0	0
1	0	1	0	0
1	0	1	1	1
1	0	0	1	0
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1