Academic Task Number: 3 Course code: INT354 Section: _____

Course title: Machine Learning-1 Maximum Marks 30

Roll No.: _____ Name: _____Registration No.: _____

1. What does the Bayesian network provides?
   a) Complete description of the domain
   b) Partial description of the domain
   c) Complete description of the problem
   d) None of the mentioned

2. Formula for Bayes theorem is _____
   a) P(A|B) = P(B | A). P(A) / P(B)
   b) P(A|B) = P(A)P(B)
   c) P(A|B) = P(B | A) P(B)
   d) P(A|B) = 1P(B)

3. Method in which the previously calculated probabilities are revised with values of new probability is called _____
   a) Revision theorem
   b) Bayes theorem
   c) Dependent theorem
   d) Updation theorem

4. Formula for conditional probability P(A|B) is _____
   a) P(A|B) = P(A∩B)/P(B)
   b) P(A|B) = P(A∩B)P(A)
   c) P(A|B) = P(A)P(B)
   d) P(A|B) = P(B)P(A)

5. It is observed that 50% of mails are spam. There is a software that filters spam mail before reaching the inbox. It accuracy for detecting a spam mail is 99% and chances of tagging a non-spam mail as spam mail is 5%. If a certain mail is tagged as spam find the probability that it is not a spam mail.
   a) 5.3% approx.
   b) 4.8% approx.
   c) 3.9% approx.
   d) 5.7% approx.

Academic Task Number: 3 Course code: INT354 Section: _____

Course title: Machine Learning-1 Maximum Marks 30

Roll No.: _____ Name: _____Registration No.: _____

6. Machine learning algorithms build a model based on sample data, known as .....
   A. Training Data
   B. Transfer Data
   C. Data Training
   D. None of the above

7. Logistic regression is a ..... regression technique that is used to model data having a ..... outcome.
   A. Linear, binary
   B. Linear, numeric
   C. Non-linear, binary
   D. Non-linear, numeric

   If the value of k is very large in the KNN algorithm, the model will underfit **the training data.**

8. If value of k is very small in KNN algorithm, model is
   a) Underfitting
   b) Overfitting
   c) Perfect fit
   d) None of these

9. A training set is called epsilon-representative if
a) For every h, Ls(h)-Ld(h)>=epsilon
b) For every h, Ls(h)-Ld(h)<=epsilon
c) For every h, |Ls(h)-Ld(h)|>=epsilon
d) For every h, |Ls(h)-Ld(h)|<=epsilon

10. What is used to measure the uniform convergence?
    a) Rademacher complexity
    b) VC-dimension
    c) Natarajan dimension
    d) All of these

11. VC dimesion is used for
a) Finite hypothesis and multiclass classification problem
b) Infinite hypothesis and multiclass classification problem
c) Finite hypothesis and binary classification problem
d) infinite hypothesis and binary classification problem.

Academic Task Number: 3 Course code: INT354 Section: _____

Course title: Machine Learning-1 Maximum Marks 30

Roll No.: _____ Name: _____Registration No.: _____

12. Natarajan dimension is the generalization of
a) Redemacher complexity
b) VC-dimensiom
c) Non-uniform learnability
d) Consistency Learnability

13. According to no free lunch theorem:
a) One classifier can be prefer over another without prior knowledge
b) All classifier perform equally if performance is taken average overall objective functions.
c) One feature can be prefer over another without prior knowledge
d) All classifier do not perform equally if performance is taken average overall objective functions

14. Choose the correct statement:
a) As the hypothesis class increases, approximation error increases and estimation error decreases.
b) As the hypothesis class increases, approximation error decreases and estimation error increases.
c) As the hypothesis class decreases, approximation error increases and estimation error decreases.
d) As the hypothesis class decreases, approximation error decreases and estimation error increases.

15. Consider the following confusion matrix. What is the precision of the model?

predicted→ real ↓	Class_pos	Class_neg
Class_pos	114	86
Class_neg	7	93

a) 0.57
b) 0.75
c) 0.94
d) 0.4

16. Complete the given statement of code snippet if the 70% of the data is given for training the model.

X_train,X_test,y_train,y_test=train_test_split(X,y,_____,random_state=0)
A. test_size=0.3
B. test_size=0.7
C. test_shape=0.3
D. None of these

Academic Task Number: 3 Course code: INT354 Section: _____

Course title: Machine Learning-1 Maximum Marks 30

Roll No.: _____ Name: _____Registration No.: _____

17. Consider the given dataset:

Swim	Wings	Green Color	Dangerous Teeth	Animal Typ
50/500	500/500	400/500	0	Parrot
450/500	0	0	500/500	Dog
500/500	0	100/500	50/500	Fish

How many total numbers of examples are present in the dataset?

A. 1000
B. 500
C. 1500
D. can't be determined

18. In Bayes theorem, the previous probabilities that are updated by using new available information is called as:

a) prior probabilities
b) independent probabilities
c) dependent probabilities
d) posterior probabilities

19. To predict the weather of upcoming week' is an example of which of the following?

A. Supervised Machine Learning: classification
B. S**upervised Machine learning: regression**
C. Unsupervised Machine Learning
D. Reinforcement learning

20. Choose the correct statement in terms of handling the overfitting?

I. Increase the dimensionality of data
II. Decrease the dimensionality of data
III. Use regularization method
IV. Use kernel approach

A. I and III
B. I and II
C. II and IV
D. II and III

Academic Task Number: 3 Course code: INT354 Section: _____

Course title: Machine Learning-1 Maximum Marks 30

Roll No.: _____ Name: _____Registration No.: _____

21. Choose the correct statement/statements:

S1: Regularization is one approach to tackle the problem of overfitting

S2: The difference between ridge and lasso regression is that lasso tends to make coefficients to absolute zero as compared to Ridge which never sets the value of the coefficient to absolute zero

a) S1 is true and S2 is true
b) S1 is true and S2 is false
c) S1 is false and S2 is true
d) S1 is false and S2 is false

22. Choose the correct statement out of the given statements:

S1: polynomial regression analysis is used to represent a non-linear relationship between dependent and independent variables.

S2: polynomial regression is a variant of the multiple linear regression model, except that the best fit line is curved rather than straight.

a) S1 is true and S2 is true
b) S1 is true and S2 is false
c) S1 is false and S2 is true
d) S1 is false and S2 is false

23. Choose the correct statement/statements:

S1: Every very decision tree has high variance

S2: A Random Forest is an ensemble technique capable of performing both regression and classification tasks with the use of multiple decision trees

S3: In the case of a regression problem, to calculate the final output in Decision trees we use majority voting.

a) S1 is true and S2 is true and S3 is true
b) S1 is true and S2 is false and S3 is false
c) S1 is true and S2 is true and S3 is false
d) S1 is false and S2 is false and S3 is true

24. Which of the following regression model uses Sigmoid activation function ?

a) Linear Regression
b) Polynomial regression
c) Multiple regression
d) Logistic regression

25. *RANSAC* is a a non-deterministic iterative algorithm that estimates the parameter of a _____ learning algorithm from a dataset that contains outliers.

a) Unsupervised
b) Supervised
c) Reinforcement
d) None of the given options

Academic Task Number: 3 Course code: INT354 Section: _____

Course title: Machine Learning-1 Maximum Marks 30

Roll No.: _____ Name: _____Registration No.: _____

26. Choose the correct statement/statements:

S1: RANSAC model can estimate the parameters with a high degree of accuracy even when a significant number of outliers are present in the data set.

S2: In RANSAC model number of iterations increases logarithmically with outlier percentage

a) S1 is true and S2 is true
b) S1 is true and S2 is false
c) S1 is false and S2 is true
d) S1 is false and S2 is false

27. Choose the correct statement/statements:

S1: If there are two or more highly collinear variables then Ridge regression selects one of them randomly which is not good for the interpretation of our model.

S2: Lasso regression decreases the complexity of a model but does not reduce the number of independent variables since it never leads to a coefficient being zero rather only minimizes it.

a) S1 is true and S2 is true
b) S1 is true and S2 is false
c) S1 is false and S2 is true
d) S1 is false and S2 is false

28. To plot the scatterplot matrix(for EDA), we will use the pairplot function from the _____ library.

a) Numpy
b) Pandas
c) Seaborn
d) Matplotlib

29. Train-Test split is available in which package of sklearn?

A. preprocessing
B. train-test
C. model_selection
D. metrics

30. Choose the correct statement in terms of handling the overfitting?

I. Increase the dimensionality of data
II. Decrease the dimensionality of data
III. Use regularization method
IV. Use kernel approach

a) I and III
b) I and II
c) II and IV
d) II and III