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

Course title: Machine Learning-1 Maximum Marks 30

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

1. Regarding bias and variance, which of the following statements are true?

A. Models which overfit have a high bias.

B. Models which overfit have a low bias.

C. Models which underfit have a high variance.

D. None of the mentioned

2. In a particular pain clinic, 10% of patients are prescribed narcotic pain killers. Overall, five percent of the clinic's patients are addicted to narcotics (including pain killers and illegal substances). Out of all the people prescribed pain pills, 8% are addicts. If a patient is an addict, what is the probability that they will be prescribed pain pills?

a) 0.16
b) 0.008
c) 0.08
d) 0.01

3. Choose the correct statement/statements:

S1: The correlation matrix is a square matrix that contains the Pearson product-moment correlation coefficients (often abbreviated as Pearson's r), which measure the linear dependence between pairs of features.

S2:The correlation coefficients are bounded to the **range 0 and 1**

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

4. To represent perfect positive correlation the Pearson coefficient in Correlation analysis should be_____

a) 0
b) -1
c) 1
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.: _____

5. *Which one is true?*
   (A) Ridge regression decreases the complexity of a model but does not reduce the number of variables since it never leads to a coefficient been zero rather only minimizes it
   (B) Lasso regression is not good for feature reduction
   (C) As the regularization parameter increases, the value of the coefficient tends towards zero. This leads to both low variance (as some coefficient leads to negligible effect on prediction) and low bias (minimization of coefficient reduces the dependency of prediction on a particular variable)
   a) Only A and B
   b) Only A, B and C
   c) Only A and C
   d) All A, B and C

6. The strength (degree) of the correlation between a set of independent variables X and a dependent variable Y is measured by————
   A : **Coefficient of Correlation**
   B : Coefficient of Determination
   C : Standard error of estimate
   D : Probability

7. 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.

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

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

Course title: Machine Learning-1 Maximum Marks 30

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

9. 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) 3.9% approx.
   c) 5.7% approx.
   d) 4.8% approx.

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

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

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

13. Natarajan dimension is the generalization of
a) Rademacher complexity
b) Non-uniform learnability
c) VC-dimensioN
d) Consistency Learnability

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

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

Course title: Machine Learning-1 Maximum Marks 30

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

15. Choose the correct statement:

a) As the hypothesis class increases, approximation error decreases and estimation error increases.

b) As the hypothesis class increases, approximation error increases and estimation error decreases.

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.

16. 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) 0.75

b) 0.57

c) 0.94

d) 0.4

17. Complete the given statement of code snippet if the 90% 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=0.1

B. test_size=0.2

C. test_shape=0.3

D. None of these

18. *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.: _____

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

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

20. To predict the "stock market analysis" is an example of which of the following?

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

21. 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 false
b) S1 is false and S2 is true
c) S1 is true and S2 is true
d) S1 is false and S2 is false

22. 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. II and III
C. I and II
D. II and IV

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

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

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

Course title: Machine Learning-1 Maximum Marks 30

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

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

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

25. 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) II and III
c) I and II
d) II and IV

26. 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. 1500
B. 1000
C. 500
D. can't be determined

27. Choose the correct statement/statements:

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

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

Course title: Machine Learning-1 Maximum Marks 30

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

d) S1 is false and S2 is false

28. Choose the correct statement/statements:

S1: Every very decision tree has low 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 false and S2 is true and S3 is false
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

29. 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

30. What does the Bayesian network provides?

a) Partial description of the domain
b) Complete description of the problem
c) Complete description of the domain
d) None of the mentioned