1. Regarding bias and variance, which of the following statements are true?
2. Models which overfit have a high bias.
3. Models which overfit have a low bias.
4. Models which underfit have a high variance.
5. None of the mentioned
6. 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?
7. 0.16
8. 0.008
9. 0.08
10. 0.01
11. 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

1. S1 is true and S2 is true
2. S1 is true and S2 is false
3. S1 is false and S2 is true
4. S1 is false and S2 is false
5. To represent perfect positive correlation the Pearson coefficient in Correlation analysis should be\_\_\_\_\_
6. 0
7. -1
8. 1
9. None of the given options
10. *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)
11. Only A and B
12. Only A, B and C
13. Only A and C
14. All A, B and C
15. 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

1. Choose the correct statement:
2. As the hypothesis class increases, approximation error increases and estimation error decreases.
3. As the hypothesis class increases, approximation error decreases and estimation error increases.
4. As the hypothesis class decreases, approximation error increases and estimation error decreases.
5. As the hypothesis class decreases, approximation error decreases and estimation error increases.
6. Formula for Bayes theorem is \_\_\_\_\_\_\_\_  
   a) P(A|B) = P(A)P(B)
7. P(A|B) = P(B│A) P(A)P(B)
8. P(A|B) = P(B│A) P(B)
9. P(A|B) = 1P(B)
10. It is observed that 50% of mails are spam. There is a software that filters spam mail before reaching the inbox. Its accuracy for detecting 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.
11. 5.3% approx.
12. 3.9% approx.
13. 5.7% approx.
14. 4.8% approx.
15. Machine learning algorithms evaluate a model based on sample data, known as .................

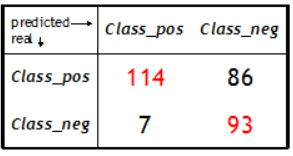
A. Testing Data

B. Transfer Data

C. Data Training

D. None of the above

1. If value of k is very large in KNN algorithm, model is
2. Underfitting
3. Overfitting
4. Perfect fit
5. None of these
6. What is used to measure the uniform convergence?
7. VC-dimension
8. Natarajan dimension
9. All of these
10. Rademacher complexity
11. Natarajan dimension is the generalization of
12. Rademacher complexity
13. Non-uniform learnability
14. VC-dimensiom
15. Consistency Learnability
16. According to no free lunch theorem:
17. One classifier can be prefer over another without prior knowledge
18. One feature can be prefer over another without prior knowledge
19. All classifier do not perform equally if performance is taken average overall objective functions
20. All classifier perform equally if performance is taken average overall objective functions.
21. Choose the correct statement:
22. As the hypothesis class increases, approximation error decreases and estimation error increases.
23. As the hypothesis class increases, approximation error increases and estimation error decreases.
24. As the hypothesis class decreases, approximation error increases and estimation error decreases.
25. As the hypothesis class decreases, approximation error decreases and estimation error increases.
26. Consider the following confusion matrix. What is the precision of the model?



1. 0.75
2. 0.57
3. 0.94
4. 0.4
5. 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)

1. test\_size=0.1
2. test\_size=0.2
3. test\_shape=0.3
4. None of these
5. *RANSAC* is a a non-deterministic iterative algorithm that estimates the parameter of a \_\_\_\_\_\_\_\_ learning algorithm from a dataset that contains outliers.
6. Unsupervised
7. Supervised
8. Reinforcement
9. None of the given options
10. In Bayes theorem, the previous probabilities that are updated by using new available information is called as:
11. prior probabilities
12. posterior probabilities
13. independent probabilities
14. dependent probabilities
15. To predict the “stock market analysis” is an example of which of the following?
16. Supervised Machine learning: regression
17. Supervised Machine Learning: classification
18. Unsupervised Machine Learning
19. Reinforcement learning
20. 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.

1. S1 is true and S2 is false
2. S1 is false and S2 is true
3. S1 is true and S2 is true
4. S1 is false and S2 is false
5. Choose the correct statement in terms of handling the overfitting?
6. Increase the dimensionality of data
7. Decrease the dimensionality of data
8. Use regularization method
9. Use kernel approach
10. I and III
11. II and III
12. I and II
13. II and IV
14. Which of the following regression model uses Sigmoid activation function ?
15. Linear Regression
16. Polynomial regression
17. Multiple regression
18. Logistic regression
19. To plot the scatterplot matrix(for EDA), we will use the Heatmap function from the \_\_\_\_\_\_\_\_ library.
20. Numpy
21. Pandas
22. Seaborn
23. Matplotlib
24. Choose the correct statement in terms of handling the overfitting?
25. Increase the dimensionality of data
26. Decrease the dimensionality of data
27. Use regularization method
28. Use kernel approach
29. I and III
30. II and III
31. I and II
32. II and IV
33. Consider the given dataset:



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

1. 1500
2. 1000
3. 500
4. can’t be determined
5. 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

1. S1 is true and S2 is true
2. S1 is true and S2 is false
3. S1 is false and S2 is true
4. S1 is false and S2 is false
5. 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.

1. S1 is false and S2 is true and S3 is false
2. S1 is true and S2 is false and S3 is false
3. S1 is true and S2 is true and S3 is false
4. S1 is false and S2 is false and S3 is true
5. A training set is called epsilon-representative ifA training set is called epsilon-representative if
6. For every h, |Ls(h)-Ld(h)|<=epsilon
7. For every h, Ls(h)-Ld(h)>=epsilon
8. For every h, Ls(h)-Ld(h)<=epsilon
9. For every h, |Ls(h)-Ld(h)|>=epsilon
10. What does the Bayesian network provide?
11. Partial description of the domain
12. Complete description of the problem
13. Complete description of the domain
14. None of the mentioned