PFA_FILE_3

Q1. Among the following identify the one in which dimensionality reduction reduces ?

ANS :- Collinearity.

Q2. Which of the following machine learning algorithm is based upon the idea of bagging ?

ANS :- Random Forest.

Q3. Choose a disadvantage of decision trees among the following ?

ANS :- Decision Tree are prone to overfit.

Q4. What is the term known as on which the machine learning algorithms build a model based on sample data?

ANS :- Training data.

Q5. Which of the following machine learning techniques helps in detecting the outliers in data ?

ANS :- Anamoly detection.

Q6. Identify the incorrect numerical functions in the various function representation of machine learning ?

ANS :- Case based.

Q7. Analysis of ML algorithm needs ?

ANS :- Both a and b.

Q8. Identify the difficulties with the k-nearest neighbor algorithm ?

ANS :- Both a and b.

Q9. The total types of the layer in radial basis function neural networks is _____ ?

ANS :- 3.

Q10. Which of the following is not a supervised learning ?

ANS :- PCA.

Q11. What is unsupervised learning ?

ANS :- Neither feature nor number of groups is known.

Q12. Which of the following is not a machine learning algorithm ?

ANS :- SVG.

Q13. _____ is the scenario when the model fails to decipher the underlying trend in the input data ?

ANS :- Underfitting.

Q14. Real-Time decisions, Game AI, Learning Tasks, Skill acquisition, and Robot Navigation are applications of ..... ?

ANS :- Reinforcement learning.

Q15. What is called the average squared difference between classifier predicted output and actual output ?

ANS :- Mean squared error.

Q16. Logistic regression is a ..... regression technique that is used to model data having a ..... outcome ?

ANS :- Nonlinear, binary.

Q17. You are given reviews of few netflix series marked as positive, negative and neutral. Classifying reviews of a new netflix series is an example of ?

ANS :- supervised learning.

Q18. Following is powerful distance metrics used by Geometric model ?

ANS :- euclidean distance.

Q19. Which of the following techniques would perform better for reducing dimensions of a data set?

ANS :- removing columns which have too many missing values.

Q20. Supervised learning and unsupervised clustering both require which is correct according to the statement ?

ANS :- input attribute.

Q21. What is the meaning of hard margin in SVM ?

ANS :- SVM allows very low error in classification. SVM allows high amount of error in classification

Q22. Increase in which of the following hyper parameter results into overfit in Random forest? (1). Number of Trees. (2). Depth of Tree, (3). Learning Rate ?

ANS :- Only 2.

Q23. Below are the 8 actual values of target variable in the train file: [0,0,0, 0, 1, 1,1,1,1,1], What is the entropy of the target variable ?

ANS :- -(6/10 log(6/10) + 4/10 log(4/10)).

Q24. Lasso can be interpreted as least-squares linear regression where ?

ANS :- weights are regularized with the l1 norm.

Q25. Consider the problem of binary classification. Assume I trained a model on a linearly separable training set, and now I have a new labeled data point that the model properly categorized and is far away from the decision border. In which instances is the learnt decision boundary likely to change if I now add this additional point to my previous training set and re-train? When the training model is ?

ANS :- Logistic regression and Gaussian discriminant analysis.

Q26. Assume you've discovered multi-collinear features. Which of the following actions do you intend to take next? (1). Both collinear variables should be removed. (2). Instead of deleting both variables, we can simply delete one. (3). Removing correlated variables may result in information loss. We may utilize penalized regression models such as ridge or lasso regression to keep such variables ?

ANS :- Either 2 or 3.

Q27. A least squares regression study of weight (y) and height (x) yielded the following least squares line: y = 120 + 5x. This means that if the height is increased by one inch, the weight should increase by what amount ?

ANS :- increase by 1 pound.

Q28. The line described by the linear regression equation (OLS) attempts to ____ ?

ANS :- Pass through as many points as possible.

Q29. For two real-valued attributes, the correlation coefficient is 0.85. What does this value indicate ?

ANS :- The attributes are not linearly related.

Q30. Which neural network architecture would be most suited to handle an image identification problem (recognizing a dog in a photo) ?

ANS :- Convolutional Neural Network.