QUIZ TOPIC - CLASSIFICATION

1. **Which of the following metrics are used to evaluate classification models?**

- ○ A. Area under the ROC curve
- ○ B. F1 score
- ○ C. Confusion matrix
- ◉ D. All of the above ✔

2. **Which one is a classification algorithm?**

- ◉ A. Logistic regression ✔
- ○ B. Linear regression
- ○ C. Polynomial regression
- ○ D. None

3. **Classification is-**

- ○ A. Unsupervised learning
- ○ B. Reinforcement learning
- ◉ C. Supervised learning ✔
- ○ D. None

4. **You have a dataset of different flowers containing their petal lengths and color. Your model has to predict the type of flower for given petal lengths and color. This is a-**

- ○ A. Regression task
- ◉ B. Classification task ✔
- ○ C. Clustering task
- ○ D. None

5. **A classifier-**

- ○ A. Inputs a vector of continuous values and outputs a single discrete value
- ○ B. Inputs a vector of discrete values and outputs a single discrete value
- ◉ C. Both A and B ✔
- ○ D. None

6. **Classification is appropriate when you-**

- ○ A. Try to predict a continuous valued output

- ◉ B. Try to predict a class or discrete output ✔
- ○ C. Both A and B for different contexts
- ○ D. None

7. **With the help of a confusion matrix, we can compute-**

- ○ A. Recall
- ○ B. Precision
- ○ C. Accuracy
- ◉ D. All of the above ✔

8. **What does recall refer to in classification?**

- ◉ A. The proportion of all the relevant data points ✔
- ○ B. The proportion of only the correct data points
- ○ C. The proportion of all data points
- ○ D. None

9. **False negatives are-**

- ◉ A. Predicted negatives that are actually positives ✔
- ○ B. Predicted positives that are actually negatives
- ○ C. Predicted negatives that are actually negatives
- ○ D. Predicted positives that are actually positives

10. **Suppose your classification model predicted true for a class which actual value was false. Then this is a-**

- ◉ A. False positive ✔
- ○ B. False negative
- ○ C. True positive
- ○ D. True negative

11. **The false negative value is 5 and the true positive value is 20. What will be the value of recall-**

- ○ A. 0.2
- ○ B. 0.6
- ◉ C. 0.8 ✔
- ○ D. 0.3

12. **The true positive value is 10 and the false positive value is 15. Calculate the value of precision-**

- A. 0.6
- B. 0.4 ✔
- C. 0.5
- D. None

13. **If the precision is 0.6 and the recall value is 0.4. What will be the f measure?**

- A. 0.5 ✔
- B. 0.6
- C. 0.4
- D. 0.3

14. **Which one is a different algorithm?**

- A. Logistic Regression ✔
- B. Support Vector Regression
- C. Polynomial Regression
- D. None

15. **What is a support vector?**

- A. The distance between any two data points
- B. The average distance between all the data points
- C. The distance between two boundary data points
- D. The minimum distance between any two data points

16. **Which of the following is a lazy learning algorithm?**

- A. SVM
- B. KNN ✔
- C. Decision tree
- D. All of the above

17. **Which of the following is not a lazy learning algorithm?**

- A. SVM
- B. Decision tree
- C. Random forest
- D. All of the above ✔

18. **What is the most widely used distance metric in KNN?**

- (●) A. Euclidean distance ✔
- ( ) B. Manhattan distance
- ( ) C. Perpendicular distance
- ( ) D. All of the above

19. **Which of the following is the best algorithm for text classification?**

- ( ) A. KNN
- ( ) B. Decision tree
- ( ) C. Random forest
- (●) D. Naive Bayes ✔

20. **What does k stand for in the KNN algorithm?**

- (●) A. Number of neighbors ✔
- ( ) B. Number of output classes
- ( ) C. Number of input features
- ( ) D. None

21. **Support Vector Machine is-**

- (●) A. a discriminative classifier ✔
- ( ) B. a lazy learning classifier
- ( ) C. a probabilistic classifier
- ( ) D. None

22. **What are hyperplanes?**

- (●) A. Decision boundaries ✔
- ( ) B. Decision functions
- ( ) C. Mapping functions
- ( ) D. None

23. **What is a kernel?**

- ( ) A. A function that calculates the distance of two boundary data points
- (●) B. A function that maps the value from one dimension to the other ✔
- ( ) C. A function that predicts the output value of a regression
- ( ) D. None

24. **Which of the following is not a kernel?**

- ( ) A. Polynomial Kernel

- ○ B. Gaussian Kernel
- ○ C. Sigmoid Kernel
- ◉ D. None ✔

25. **Why Naive Bayes is called naive?**

- ◉ A. Because its assumption may or may not true ✔
- ○ B. Because it's a bad classifier
- ○ C. The accuracy is very poor
- ○ D. All of the above

26. **For two events A and B, the Bayes theorem will be-**

- ○ A. P(A | B) = P(B) * P(B | A) / P(A)
- ◉ B. P(A | B) = P(A) * P(B | A) / P(B) ✔
- ○ C. P(A | B) = P(B) * P(A | B) / P(A)
- ○ D. P(A | B) = P(A) * P(A | B) / P(B)

27. **How does a decision tree work?**

- ○ A. Minimizes the information gain and maximizes the entropy
- ◉ B. Maximizes the information gain and minimizes the entropy ✔
- ○ C. Minimizes the information gain and minimizes the entropy
- ○ D. Maximizes the information gain and maximizes the entropy

28. **Suppose you have a dataset that is randomly distributed. What will be the best algorithm for that dataset?**

- ○ A. Support vector machine ✖
- ○ B. Naive Bayes
- ○ C. K nearest neighbors
- ◉ D. Decision tree ✔

29. **Which pair of the algorithms are similar in operation?**

- ○ A. SVM and KNN
- ◉ B. Decision tree and Random forest ✔
- ○ C. SVM and Naive Bayes ✖
- ○ D. All of the above ✖

30. **Which metric is not used for evaluating classification models?**

- ○ A. AUC ROC score

B. Accuracy

C. Mean absolute error ✔

D. Precision