**MIDTERM QUESTIONS**

1. Which type of machine learning has a reward system?

a) Supervised Learning

**b) Reinforcement Learning**

c) Unsupervised Learning

d) Deep Neural Networks

2. Which is not a step in preprocessing?

a) Feature extraction

b) Sampling

**c) Clustering**

d) Dimensionality Reduction

3. For predicting continuous values we use

a) Classification

**b) Regression**

c) Dimensionality Reduction

d) Clustering

4. Which one of the following is a step in Learning?

a) EDA

b) Evaluation

c) Preprocessing

**d) Performance Metrics**

5. Choose the right order of predictive modelling

**a) Preprocessing, Learning, Evaluation, Prediction**

b) Learning, Evaluation, Prediction

c) Prediction, Learning, Evaluation

d) Preprocessing, Evaluation, Learning, Prediction

6. Perceptron works best on

a) All kinds of data

b) Non-linearly separable data

**c)** **Linearly separable data**

d) Colinear classes

7. What are the three dimensions used to represent a color image in Python?

**a) Image height, Image width and color channel**

b) Red, Blue and Green channels

c) Hue, Saturation and Lightness

d) None of the above

8. In Adaline, weights are updates based on

**a) Linear activation function**

b) Sigmoid function

c) Unit step function

d) Relu

9. What is delta (error) in perceptron model of neuron?  
a)difference between desired & target output  
**b)** **error due to environmental condition**  
c) can be both due to difference in target output or environmental condition  
d) none of the mentioned

10. what is the another name of weight update rule in adaline model based on its functionality?  
a) LMS error learning law  
b) gradient descent algorithm  
**c) both LMS error & gradient descent learning law**  
d) Sum of squared errors

11. The sigmoid functions follows the shape

a) Bell curve

b) straight line

c) parallel to one of the axis

**d) S curve**

12. The training samples closest to the separating hyperplane are called as

a) margin

**b) Support Vectors**

c) decision boundary

d) Training vectors

13. One the most widely used kernel function in SVM is

a) Linear

b) polynomial

**c)** **RBF**

d) Non-linear

14. Which of the following is not a splitting criteria that’s commonly used in Binary Decision Trees?

**a) Kernel**

b) Gini impurity

c) entropy

d) classification error

15. How many hyperplanes does a n-dimensional space have?

a) n

**b) n-1**

c) n/2

d) n^1/2

16. dropna(axis = 0) removes

**a) rows with missing values**

b) columns with missing values

c) All NaN values

d) All Null values

17. The method used to learn the parameters from the training data is

a) Transform

b) predict

**c)** **fit**

d) fit\_transform

18. Pick the algorithm that doesn’t worry about feature scaling

a) Support Vector Machines

b) LSTM Networks

c) KNN

**d) Random Forest**

19. Which of the following is not a common method to reduce generalization is

a) Having more Training data

**b) Choose model with high variance**

c) Choose simpler model

d) Reduce the dimensionality of data

20. Which of the following algorithm are not an example of ensemble learning algorithm?   
a) Adaboost  
b) Random forest  
c) Gradient Boosting  
**d) Decision Trees**

21. \_\_\_\_\_\_ is used for non-linear dimension reduction

**a) Kernel Principal Component Analysis**

b) Linear Discriminant Analysis

c) Principal Component Analysis

d) Independent Component Analysis

22. Identify the step that does not belong to extracting the principal components

a) Standardizing the data

**b) Sorting the eigen values by increasing order to rank the eigenvectors.**

c) Obtaining the eigenvalues and eigenvectors of the covariance matrix.

d) Constructing the covariance matrix.

23. The algorithm that finds the orthogonal component axes of maximum variance in a dataset is

a) FastICA

b) LDA

**c)** **PCA**

d) None of the above

24. LDA is a technique used for   
a) Unsupervised data compression  
b) Non Linear Dimensionality reduction  
c) Co-Linear Dimensionality Reduction  
**d) Supervised Dimensionality Reduction**

25.To use Function plot\_decision\_regions we need to import

a) Seaborn

**b) mlxtend.plotting**

c) matplotlib

d) pyplot

26. Which of the following is a categorical outcome?  
a) RMSE  
b) Rsquared  
**c) Accuracy**  
d) Least squared error

27. To balance the dataset before training the tree to prevent

biasing, one should

**a) sample an equal number of samples from each class**

b) Check ratio of samples to number of features

c) Tune min\_samples\_leaf to control the number of samples at leaf

node.

d) perform PCA

28.Which of the following correctly describes the relationship

between the complexity and the bias and variance terms

a) As the complexity decreases, the bias term decreases while

the variance term increases.

**b) As the complexity decreases, bias remains the same and the variance**

**terms decreases.**

c) As the complexity increases, the bias term increases while the

variance term decreases.

d) As the complexity increases, the bias term decreases while the

variance term increases.

29. \_\_\_\_\_helps us obtain reliable estimates of the model's generalization performance, how well the model performs on unseen data

a) Hold out cross validation

**b) Grid Search CV**

c) K Fold CV

d) LOOCV

30. For which of the following reasons would you want to use a hold-out set for the very end?

**a) You want to be absolutely certain about your model's ability to generalize to unseen data.**  
b) You want to tune the hyperparameters of your model.  
c) You want to maximize the amount of training data used.  
d) None of the above