Day-12 Quiz-DataScience-Training

Welcome to the Python Programming Quiz! This quiz tests your knowledge of daily learnings. Please read the instructions carefully before starting the quiz.

Instructions and Rules

- **Time Limit:** You have 20 minutes to complete the quiz.
- **Number of Questions:** The quiz consists of 20 multiple-choice questions.
- **Scoring:** Each correct answer is worth 1 point. There is no negative marking for incorrect answers.
- **Single Attempt:** You are allowed only one attempt to complete the quiz.
- **Required Fields:** All questions are mandatory. You must answer each question to submit the quiz.
- **Resources:** This is a closed-book quiz. Do not use any external resources, including books, notes, or the internet.
- **Honesty:** Please answer the questions honestly and to the best of your ability. Cheating or dishonesty will result in disqualification.
- **Environment:** Ensure you are in a quiet environment where you can concentrate without interruptions.
- **Technical Issues:** In case of technical issues, please contact the quiz administrator immediately.
- **Retakes:** There are no retake opportunities for this quiz. Ensure you are prepared before starting.

Good luck, and do your best!

* Indicates required question

1. Email *

_____

2. **1. What is Ensemble Learning?** *

Mark only one oval.

○ a) A single machine learning model

○ b) A technique that uses multiple models to improve performance

○ c) A data preprocessing method

○ d) A type of neural network

3. **2. Which of the following is NOT an ensemble method?** *

Mark only one oval.

○ a) Bagging

○ b) Boosting

○ c) Stacking

○ d) Clustering

4. **3. What is the primary advantage of using ensemble methods?** *

Mark only one oval.

- ◯ a) Reducing data size
- ◯ b) Enhancing the interpretability of models
- ◯ c) Improving model accuracy and robustness
- ◯ d) Simplifying the model

5. **4. Which ensemble method involves training multiple models on different subsets of the data and averaging their** * **predictions?**

Mark only one oval.

- ◯ a) Boosting
- ◯ b) Bagging
- ◯ c) Stacking
- ◯ d) Blending

6. **5. Which ensemble method focuses on training models sequentially, where each model tries to correct the errors** * **of the previous one?**

Mark only one oval.

- a) Bagging
- b) Boosting
- c) Stacking
- d) Voting

7. **6. What is the primary purpose of Principal Component Analysis (PCA)?** *

Mark only one oval.

- a) Data classification
- b) Data clustering
- c) Dimensionality reduction
- d) Data normalization

8.   **7. Which of the following best describes the first principal component in PCA?** *

Mark only one oval.

- ( ) a) The component with the least variance
- ( ) b) The component orthogonal to all others
- ( ) d) The component with the smallest eigenvalue
- ( ) d) The component with the highest variance

9.   **8. What is the primary mathematical tool used in PCA to find the principal components?** *

Mark only one oval.

- ( ) a) Logistic regression
- ( ) b) Eigen decomposition
- ( ) c) Gradient descent
- ( ) d) K-means clustering

10. **9. Which of the following is a crucial step in preparing data for PCA?** *

Mark only one oval.

- ⬭ a) Normalizing the data
- ⬭ b) Encoding categorical variables
- ⬭ c) Applying decision trees
- ⬭ d) Removing outliers

11. **10. How do you handle missing values in the dataset before applying PCA?** *

Mark only one oval.

- ⬭ a) Ignore the missing values
- ⬭ b) Replace them with zeros
- ⬭ c) Impute missing values with mean or median
- ⬭ d) Remove entire dataset

12. **11. What is the purpose of using a standard scaler before applying PCA?** *

Mark only one oval.

- a) To increase the variance of the data
- b) To standardize features by removing the mean and scaling to unit variance
- c) To normalize the data to the range [0, 1]
- d) To reduce the dimensionality of the data

13. **12. Which of the following statements is true about feature scaling in ensemble methods?** *

Mark only one oval.

- a) Feature scaling is always necessary
- b) Feature scaling is never necessary
- c) Feature scaling is generally not required
- d) Feature scaling only affects the performance of decision trees

14. **13. How do you choose the number of principal components to retain in PCA?** *

Mark only one oval.

- ◯ a) By keeping all components
- ◯ b) By selecting components with the highest eigenvalues
- ◯ c) By choosing components that explain a cumulative variance above a certain threshold
- ◯ d) By random selection

15. **14. What is a key step in the workflow of implementing a Random Forest model?** *

Mark only one oval.

- ◯ a) Splitting the dataset
- ◯ b) Calculating eigenvalues
- ◯ c) Performing PCA
- ◯ d) Applying LSTM

16. **15. What is the primary parameter to tune in PCA to control the number of components?** *

Mark only one oval.

- ◯ a) n_estimators
- ◯ b) max_depth
- ◯ c) n_components
- ◯ d) learning_rate

17. **16. What does the confusion matrix represent in the context of digit classification?** *

Mark only one oval.

- ◯ a) The variance of principal components
- ◯ b) The true and predicted classes of digits
- ◯ c) The clustering of digit images
- ◯ d) The dimensionality of the dataset

18. **17. Which of the following is a disadvantage of ensemble methods? ***

Mark only one oval.

○ a) Increased variance

○ b) Increased bias

○ c) Higher computational cost

○ d) Lower accuracy

19. **18. What is a common use case for the AdaBoost algorithm? ***

Mark only one oval.

○ a) Dimensionality reduction

○ b) Classification problems

○ c) Clustering

○ d) Data normalization

20. **19. Which of the following best describes the Gradient Boosting algorithm? ***

Mark only one oval.

○ a) A method that combines weak learners sequentially to create a strong learner

○ b) A method that applies gradient descent to logistic regression

○ c) A method that combines strong learners in parallel to create a weak learner

○ d) A method that only works with decision trees

21. **20. What is the primary difference between AdaBoost and Gradient Boosting? ***

Mark only one oval.

○ a) AdaBoost builds models sequentially; Gradient Boosting builds them in parallel

○ b) AdaBoost uses only decision trees; Gradient Boosting can use any model

○ b) AdaBoost is used for regression; Gradient Boosting is used for classification

○ d) AdaBoost assigns weights to misclassified samples; Gradient Boosting uses gradients