ML Intermediate Assessment

TE MARKS (/PACKAGE-COST-DETAILS/REGENERATE-MARKS?STUDENTID=20453&ASSESSMENTID=215557)

Score Obtained: 47/50 (94%)

RETAKE (/PACKAGE-COST-DETAILS/#) VIEW REPORT

Machine learning is an application of _____

1/1 ATTEMPTED

- Block Chain
- Artificial Intelligence
- Both A and B
- None of these

2. What is the main objective of Machine Learning?

1/1 ATTEMPTED

- Creating intelligent machines with emotions
- Writing programs to perform specific tasks.
- Teaching computers to learn from data and improve performance
- Designing hardware for computational tasks.

3. In which industry is machine learning commonly used for fraud detection and risk management?

1/1 ATTEMPTED

- Agriculture

- Finance
- Entertainment
- Education

4.

An observation that lies an abnormal distance from other values in a random sample from a population is refer to as?

1/1 ATTEMPTED

- Syrical point
- External point
- Outlier
- Skew point

5.

Which of the following is not a major data analysis approaches?

1/1 ATTEMPTED

- Data Mining
- Predictive Intelligence
- Business Intelligence
- Text Analytics

6. What is the primary purpose of EDA in Data Science? 1/1 ATTEMPTED

- To make data more complex
-

To simplify complex data

To discover insights & patterns in data

To visualize data

7. In EDA, which Measure of central tendency is typically represented by the height of a box in a Box Plot?

1/1 ATTEMPTED

Mean

Median

Mode

Variance

8. Which of the following information is not given by Five Number Summary?

0/1 ATTEMPTED

Mean

Median

Mode

All of the above

9. How can outliers in a numerical dataset be treated?

1/1 ATTEMPTED

Ignoring them during analysis

Replacing them with the median value

- Removing them or transforming them
- Assigning them a weight of 0

10. Which method in pandas provides a concise summary of a DataFrame, including the number of non-null values in each column? 1/1 ATTEMPTED

- df.describe()
- df.info()
- df.head()
- df.shape()

11.
How many coefficients do you need to estimate in a simple linear regression model (One independent variable)?

1/1 ATTEMPTED

- 1
- 12
- 3
- 2

12.
Linear Regression is an example of?

1/1 ATTEMPTED

- Supervised Learning
- Unsupervised Learning

- Semi-Supervised Learning
- Reinforcement Learning

13.

For what Polynomial Regression is used?

1/1 ATTEMPTED

- Find the best linear line
- Handle with non-linear and separable data
- Handle linear and separable data
- Classify binary data

14.

_____loss is mostly used in SVM, this is used in the combination of the activation function in the last layer. We use this loss to classify whether an email is a spam or not.

1/1 ATTEMPTED

- Hinge loss
- entropy loss
- cross-entropy loss
- MAE loss

15.

What is the name of a regression model in which more than one independent variable is utilized to predict the dependent variable?

1/1 ATTEMPTED

- a simple linear regression model

- a multiple regression model
- an independent model
- none of the above

16. Which of the following is a commonly used metric for evaluating the performance of a Linear Regression model? 1/1 ATTEMPTED

- Accuracy Score
- F1 Score
- Mean Squared Error (MSE)
- Confusion Matrix

17. What is a key assumption of Linear Regression? 1/1 ATTEMPTED

- The data must have a normal distribution
- The relationship between variables is linear
- Outliers have no impact on the model
- The number of features should be equal to the number of observations

18. What is the primary purpose of a Loss Function in Linear Regression? 1/1 ATTEMPTED

- To maximize the accuracy of predictions
- To minimize the difference between predicted and actual values

- To calculate the mean of the target variable
- To identify outliers in the data

19.

Ridge and Lasso regression are simple techniques to _____ the complexity of the model and prevent over-fitting which may result from simple linear regression.

1/1 ATTEMPTED

- Increase
- Eliminate
- Decrease
- None of the above

20.

Which of the following of the coefficients is added as the penalty term to the loss function in Lasso regression?

1/1 ATTEMPTED

- Absolute value of magnitude
- Squared magnitude
- Number of non-zero entries
- None of the above

21.

What type of penalty is used on regression weights in Ridge regression?

1/1 ATTEMPTED

- L0

- L2
- L1
- None of the above

22.

In Ridge regression, A hyper parameter is used called '_____' that controls the weighting of the penalty to the loss function.

1/1 ATTEMPTED

- Gamma
- Alpha
- Lambda
- None of above

23.

Which of the following is correct use of cross validation?

1/1 ATTEMPTED

- Selecting variables to include in a model
- Selecting parameters in prediction function
- Comparing predictors
- All of these

24.

Which of the following is a common error measure?

0/1 ATTEMPTED

- Median absolute deviation

- Sensitivity
- Specificity
- All of the mentioned

25.
For Ridge Regression, if the regularization parameter = 0, what does it mean?

1/1 ATTEMPTED

- Large coefficients are not penalized
- Overfitting problems are not accounted for
- The loss function is as same as the ordinary least square loss function
- All of the above

26.
For Lasso Regression, if the regularization parameter = 0, what does it mean?

1/1 ATTEMPTED

- The loss function is as same as the ordinary least square loss function
- Can be used to select important features of a dataset
- Shrinks the coefficients of less important features to exactly 0
- All of the above

27. With Lasso Regression the influence of the hyper parameter lambda, as lambda tends to zero the solution approaches to

1/1 ATTEMPTED

- Zero

- One
- Linear Regression
- Infinity

28. What is the main purpose of Ridge and Lasso regularization in linear regression? 1/1 ATTEMPTED

- To increase the complexity of the model
- To reduce the impact of outliers
- To penalize large coefficients and prevent overfitting
- To simplify the model by removing unnecessary features

29.
Logistic regression assumes a: 1/1 ATTEMPTED

- Linear relationship between continuous predictor variables and the outcome variable.
- Linear relationship between continuous predictor variables and the logit of the outcome variable.
- Linear relationship between continuous predictor variables.
- Linear relationship between observations.

30.
Formula - TP / (TP + FN): Which of these is being represented by formula? 1/1 ATTEMPTED

- Accuracy

- Recall
- harmonic mean
- Classification Rate

31.

_____ uses Harmonic Mean in place of Arithmetic Mean as it punishes the extreme values more.

1/1 ATTEMPTED

- S-measure
- F-measure
- T-measure
- None of these

32.

Decision Tree is a display of an algorithm?

1/1 ATTEMPTED

- True
- False
- -
- -

33.

Boosting is a Sequential step process and works on weighted majority vote approach?

1/1 ATTEMPTED

- Yes

- No
- Can't Say
- None of these

34.

Adaboost selects a training subset randomly?

0/1 ATTEMPTED

- FALSE
- TRUE
- -
- -

35.

_____ relies on the intuition that the best possible next model, when combined with previous models, minimizes the overall prediction error

1/1 ATTEMPTED

- Gradient Boost
- XG Boost
- Ada Boost
- All of these

36.

In the case of AdaBoost, very short decision trees were used that only had a single split, called a _____?

1/1 ATTEMPTED

optimizer step

- Decision stump
- weak learner
- bench step

37. What is bootstrap sampling primarily used for?

1/1 ATTEMPTED

- To improve data quality
- To estimate the sampling distribution of a statistic
- To create a larger dataset
- To reduce computational time

38. What assumption does bootstrap sampling make about the original dataset?

1/1 ATTEMPTED

- It assumes the data is normally distributed.
- It assumes that the sample is a representative sample of the population.
- It assumes that the observations are independent and identically distributed (i.i.d.).
- It assumes the sample size must be large.

39. What does "bagging" stand for in machine learning?

1/1 ATTEMPTED

- Boosting Aggregation
- Bootstrap Aggregating

- Bayesian Aggregating
- Binary Aggregating

40. What is the primary purpose of bagging?

1/1 ATTEMPTED

- To reduce the bias of a model
- To increase the variance of a model
- To improve the accuracy and robustness of a model
- To simplify the model structure

41. How does bagging generate different training datasets?

1/1 ATTEMPTED

- By using a different feature set for each model
- By resampling the original dataset with replacement
- By applying a transformation to the original dataset
- By splitting the dataset into equal parts

42. What is one advantage of using bagging over a single model?

ATTEMPTED

1/1

- It always guarantees higher accuracy.
- It reduces the risk of overfitting.
- It simplifies the model complexity.

- It requires less computational resources.

43. What is Random Forest primarily used for?

1/1 ATTEMPTED

- Clustering
- Dimensionality Reduction
- Classification and Regression
- Time Series Forecasting

44. What technique does Random Forest use to create diverse trees?

1/1 ATTEMPTED

- Bagging
- Boosting
- K-means clustering
- PCA (Principal Component Analysis)

45. What is the effect of increasing the number of trees in a Random Forest model?

1/1 ATTEMPTED

- It always leads to lower accuracy.
- It increases model complexity without any benefit.
- It generally improves performance but with diminishing returns.
- It makes the model interpretability easier.

46. What is the main purpose of a decision tree in machine learning?

1/1 ATTEMPTED

- To perform clustering
- To reduce dimensionality
- To make predictions based on input features
- To visualize data distributions

47. Which algorithm is commonly used to construct decision trees?

1/1 ATTEMPTED

- K-means
- ID3
- PCA
- Gradient Descent

48. What method is commonly used to estimate the coefficients in logistic regression?

1/1 ATTEMPTED

- Least Squares
- Maximum Likelihood Estimation (MLE)
- k-means clustering
- Ridge Regression

49. Which of the following techniques is used to prevent overfitting?

1/1 ATTEMPTED

- Increasing the learning rate
- Cross-validation
- Reducing the training dataset size
- Using a linear model for complex data

50. Which algorithm is primarily used for dimensionality reduction?

1/1 ATTEMPTED

- Logistic Regression
- Decision Trees
- Principal Component Analysis (PCA)
- Support Vector Machines