ML Intermediate Assessment

ATE MARKS (/PACKAGE-COST-DETAILS/REGENERATE-MARKS?STUDENTID=21950&ASSESSMENTID=227241)

Score Obtained: 46/50 (92%)

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

1. What is Machine learning? 1/1 ATTEMPTED

- [x] The autonomous acquisition of knowledge through the use of computer program
- [ ] The autonomous acquisition of knowledge through the use of manual programs
- [ ] The selective acquisition of knowledge through the use of computer programs
- [ ] The selective acquisition of knowledge through the use of manual programs

2. What is the main objective of Machine Learning? 1/1 ATTEMPTED

- [ ] Creating intelligent machines with emotions
- [ ] Writing programs to perform specific tasks.
- [x] 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.

Function to drop the rows with NAN values_____?

1/1 ATTEMPTED

- df.drop()
- df.dropna()
- df.delete
- df.deletena()

5. 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

6. What does the term "Skewness" refer to in the context of data distribution analysis in EDA?

1/1 ATTEMPTED

- The symmetry of data distribution
-

The kurtosis of data distribution

The spread of data distribution

The presence of outliers

7. What is the purpose of encoding categorical data in machine learning?

1/1 ATTEMPTED

To increase the size of the dataset.

To reduce the number of categories

To make the data human-readable

To represent categorical data in a format suitable for machine learning algorithms

8. 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

9. What is the purpose of encoding categorical data in machine learning?

1/1 ATTEMPTED

To increase the size of the dataset

To reduce the number of categories

To make the data human-readable

- To represent categorical data in a format suitable for machine learning algorithms

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.
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

13.

True-False: Linear Regression is mainly used for Regression.

1/1 ATTEMPTED

- TRUE
- FALSE
- -
- -

14.

is a widely popular concept of information theory. It is the measure of number of bits that are needed to encode certain information based on an initial hypothesis_____?

1/1 ATTEMPTED

- Mean-Squared Loss
- Cross-Entropy
- Hinge loss
- Regression loss

15.

_____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

16.

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

17.

_____is an analytical approach to Linear Regression with a Least Square Cost Function..

1/1 ATTEMPTED

- Slope Equation
- Variable equation
- Normal Equation
- none of these

18. 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

19. 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

20.
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

21.
Ridge regression uses which norm?

1/1 ATTEMPTED

- L1
- L2

- Combination of L1 and L2
- None

22.

Which one is true?

1/1 ATTEMPTED

- Ridge and Lasso regression are techniques to reduce the model complexity and prevent over-fitting which may result from simple linear regression
- Ridge regression shrinks the coefficients and it helps to reduce the model complexity and multi-collinearity.
- Lasso regression not only helps in reducing over-fitting but it can help us in feature selection
- All of the above

23.

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

1/1 ATTEMPTED

- L0
- L2
- L1
- None of the above

24.

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

0/1 ATTEMPTED

- Gamma
- Alpha

- Lambda
- None of above

25.

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

26.

Which of the following is a common error measure?

0/1 ATTEMPTED

- Median absolute deviation
- Sensitivity
- Specificity
- All of the mentioned

27.

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

28.

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

29.

What's the penalty term for the Ridge regression?

1/1 ATTEMPTED

- the square of the magnitude of the coefficients
- the square root of the magnitude of the coefficients
- the absolute sum of the coefficients
- the sum of the coefficients

30. Elastic Net uses which Norm?

1/1 ATTEMPTED

- L1
- L2
- Both L1 & L2

- None of the above

31. 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

32. What is the main purpose of cross-validation in machine learning? 1/1 ATTEMPTED

- To increase the training time of the mode
- To assess how well the model will generalize to an independent dataset
- To maximize the number of features in the model
- To ensure the model perfectly fits the training data

33.
Which of the following is used where the target variable is of categorical nature? 1/1 ATTEMPTED

- Logistic Regression
- Knime
- Keras

- Linear Regression

34.

What's the cost function of the logistic regression?

0/1 ATTEMPTED

- Sigmoid function
- Logistic Function
- both (A) and (B)
- none of these

35.

What is the purpose of performing cross-validation?

1/1 ATTEMPTED

- To assess the predictive performance of the models
- To judge how the trained model performs outside the sample on test data
- Both A and B
- None of these

36.

Decision tree can be used for _____.

1/1 ATTEMPTED

- classification
- regression
- Both

None of these

37.

Does gradient boosted trees generally perform better than random forest?

1/1 ATTEMPTED

Yes

No

Can't Say

-

38.

Below lists some heuristics for best preparing your data for AdaBoost. Which of the following is an approach?

1/1 ATTEMPTED

Quality Data

Outlier

Noisy Data

All of these

39.

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

0/1 ATTEMPTED

Gradient Boost

XG Boost

- Ada Boost
- All of these

40. 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

41. In bootstrap sampling, how is a sample generated from the original dataset?

1/1 ATTEMPTED

- By taking the mean of the original dataset
- By randomly selecting observations with replacement
- By taking the median of the original dataset
- By using a linear transformation

42. Which of the following is a limitation of bootstrap sampling?

ATTEMPTED

1/1

- It can be computationally intensive.
- It cannot be used for small sample sizes.
- It requires a normal distribution of the data.

- It always provides biased estimates.

43. 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

44. 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.

45. What is Random Forest primarily used for?

1/1 ATTEMPTED

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

46. 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.

47. What is the primary purpose of logistic regression in machine learning? 1/1 ATTEMPTED

- To perform clustering
- To predict continuous values
- To model binary outcomes
- To reduce dimensionality

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. What is the primary goal of supervised learning?

1/1 ATTEMPTED

- To discover patterns in unlabeled data
- To predict outcomes based on labeled input data
- To reduce the dimensionality of the data
- To cluster similar data points

50. 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