

A feature F1 can take a certain value: A, B, C, D, E, & F and represents the grade of students from a college.

1) Which of the following statements is true in the following case?

- A) Feature F1 is an example of nominal variable.
- B) Feature F1 is an example of ordinal variable.
- C) It doesn't belong to any of the above category.
- D) Both of these

Imagine, you are working with "ABC" and you want to develop a machine learning algorithm which predicts the number of views on the articles.

Your analysis is based on features like author name, number of articles written by the same author on ABC in the past and a few other features. Which of the following evaluation metrics would you choose in that case?

1. **Mean Square Error**
2. **Accuracy**
3. **F1 Score**

- A) Only 1
- B) Only 2
- C) Only 3
- D) 1 and 3
- E) 2 and 3
- F) 1 and 2

Which of the following statements is/are true about “Type-1” and “Type-2” errors?

1. **Type1 is known as false positive and Type2 is known as false negative.**
2. **Type1 is known as false negative and Type2 is known as false positive.**
3. **Type1 error occurs when we reject a null hypothesis when it is actually true.**

- A) Only 1
- B) Only 2
- C) Only 3
- D) 1 and 2
- E) 1 and 3
- F) 2 and 3

Adding a non-important feature to a linear regression model may result in.

1. **Increase in R-square**
2. **Decrease in R-square**

- A) Only 1 is correct
- B) Only 2 is correct
- C) Either 1 or 2
- D) None of these

Imagine, you are solving a classification problem with a highly imbalanced class. The majority class is observed 99% of times in the training data.

Your model has 99% accuracy after taking the predictions on test data. Which of the following is true in such a case?

1. Accuracy metric is not a good idea for imbalanced class problems.
2. Accuracy metric is a good idea for imbalanced class problems.
3. Precision and recall metrics are good for imbalanced class problems.
4. Precision and recall metrics aren't good for imbalanced class problems.

A) 1 and 3

B) 1 and 4

C) 2 and 3

D) 2 and 4

True-False: Linear Regression is a supervised machine learning algorithm.

A) TRUE

B) FALSE

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

A) TRUE

B) FALSE

Which of the following methods do we use to find the best fit line for data in Linear Regression?

- A) Least Square Error**
- B) Maximum Likelihood**
- C) Logarithmic Loss**
- D) Both A and B**

Which of the following evaluation metrics can be used to evaluate a model while modeling a continuous output variable?

- A) AUC-ROC**
- B) Accuracy**
- C) Recall**
- D) Mean-Squared-Error**

Which of the following is true about Residuals ?

- A) Lower is better**
- B) Higher is better**
- C) A or B depend on the situation**
- D) None of these**

Suppose that we have N independent variables (X_1, X_2, \dots, X_n) and the dependent variable is Y . Now Imagine that you are applying linear regression by fitting the best fit line using least square error on this data.

You found that the correlation coefficient for one of its variables(Say X_1) with Y is -0.95 .

Which of the following is true for X_1 ?

- A) Relation between the X_1 and Y is weak
- B) Relation between the X_1 and Y is strong
- C) Relation between the X_1 and Y is neutral
- D) Correlation can't judge the relationship

Which of the following statements is true about outliers in Linear regression?

- A) Linear regression is sensitive to outliers
- B) Linear regression is not sensitive to outliers
- C) Can't say
- D) None of these

For a given data, the Standard Deviation is 25. If 5 is added to each observation, what is the new SD of the resulting observations?