Feature Engineering quiz

7 out of 7 correct

1. Which of the following encoding techniques is used for categorical variables with no inherent ordering?

- ◯ Ordinal Encoding
- ◉ Label Encoding
- ◯ Target Guided Ordinal Encoding
- ◯ None of the above

Explanation: Label Encoding is used for categorical variables with no inherent ordering, where each unique category is assigned a unique integer value.

2. Which encoding technique is used when there is a natural ordering between the categories of a categorical variable?

- ◉ Ordinal Encoding
- ◯ Label Encoding
- ◯ Target Guided Ordinal Encoding
- ◯ None of the above

Explanation: Ordinal Encoding is used for categorical variables with a natural ordering between the categories, where the categories are assigned integer values according to their order.

3. Covariance is a measure of:

- ◯ The strength of the linear relationship between two variables

- ◯ The degree of association between two variables
- ◯ The direction of the relationship between two variables
- ◉ All of the above

Explanation: Covariance measures both the strength and direction of the linear relationship between two variables, and is a measure of the degree of association between them.

4. Correlation check is used to:

- ◯ Determine the strength of the linear relationship between two variables
- ◯ Determine the direction of the relationship between two variables
- ◯ Determine the degree of association between two variables
- ◉ All of the above

Explanation: Correlation check is used to determine both the strength and direction of the linear relationship between two variables, and is a measure of the degree of association between them.

5. A dataset contains a categorical variable with no natural ordering. Which encoding technique would be most appropriate for this variable?

- ◯ Ordinal Encoding
- ◉ Label Encoding
- ◯ Target Guided Ordinal Encoding
- ◯ Covariance

Explanation: Label Encoding is most appropriate for a categorical variable with no natural ordering, as each unique category can be assigned a unique integer value.

6. A dataset contains a categorical variable that is strongly associated with the target variable. Which encoding technique would be most appropriate for this variable?

- ◯ Ordinal Encoding
- ◯ Label Encoding
- ◉ Target Guided Ordinal Encoding
- ◯ Covariance

Explanation: Target Guided Ordinal Encoding would be most appropriate for a categorical variable that is strongly associated with the target variable, as it assigns ordinal values to categories based on the mean of the target variable for each category.

7. A data scientist is performing feature engineering on a dataset with categorical variables. Which technique can be used to check the degree of association between two categorical variables?

- ◯ Ordinal Encoding
- ◯ Label Encoding
- ◉ Correlation Check
- ◯ Target Guided Ordinal Encoding

Explanation: Correlation check can be used to determine the degree of association between two variables, whether they are continuous or categorical.

Submit