

Assignment question and answer

1 What is a parameter?

Answer A parameter is a *value that helps define or control how something works.* It's like a setting or variable that you can adjust.

2 What is correlation?

Answer Correlation is a statistical measure that describes how two variables are related to each other.

3 Define Machine Learning. What are the main components in Machine Learning?

Answer Machine Learning (ML) is a branch of artificial intelligence that enables computers to learn patterns from data and make decisions or predictions without being explicitly programmed for every task.

4 How does loss value help in determining whether the model is good or not?

Answer Loss value is one of the most important signals for determining how well a machine-learning model is learning. However, loss alone does not fully determine whether a model is “good.” Here’s how loss helps—and where it falls short.

5 What are continuous and categorical variables?

Answer A continuous variable is a numerical variable that can take *any value within a range.*

These values can be measured with increasing precision (decimals, fractions)

A categorical variable represents *distinct groups or categories* rather than measurements.

These categories may or may not have a natural order.

6 What is sklearn.preprocessing?

Answer `sklearn.preprocessing` is a module in scikit-learn (a popular Python machine-learning library) that provides tools for transforming, scaling, normalizing, and encoding data before training machine-learning models.

7 What is a Test set?

Answer A test set is a portion of your dataset that is kept aside and never used during model training, and is used only to evaluate the final performance of a machine-learning model.

8 What does negative correlation mean?

Answer A negative correlation describes a relationship between two variables in which one increases while the other decreases. In other words, they move in opposite directions.

9 What is causation ?

Answer Causation (or causal relationship) is when one variable directly affects or produces a change in another variable. In other words, changes in the cause lead to changes in the effect.

10 What is an Optimizer?

Answer An optimizer decides how the model should change its internal parameters to perform better on the task.

11.What is `sklearn.linear_model` ?

Answer In Python, `sklearn.linear_model` is a module in the scikit-learn library that provides a variety of linear models for regression and classification tasks. These models assume a linear relationship between input features (`X`) and the target variable (`y`).

12 What are continuous and categorical variables?

Answer Continuous and categorical variables are two main types of variables used in statistics and machine learning. Understanding the difference is crucial for choosing the right analysis or model.

13 What is feature scaling? How does it help in Machine Learning?

Answer Feature scaling is a data preprocessing technique used in machine learning to standardize the range of independent variables (features) so that they contribute equally to the model. In simpler terms, it's about bringing all features onto a similar scale

14 How do we perform scaling in Python?

Answer Scaling features in Python is very straightforward, especially using the `scikit-learn` library, which provides built-in tools for most scaling methods

15 Explain data encoding?

Answer Data encoding is a preprocessing technique in machine learning where categorical or textual data is converted into numerical form so that algorithms, which typically work with numbers, can understand and process it. Most ML models cannot work directly with text or categorical labels, so encoding is essential.