1. What is ML?

Machine Learning (ML) is a subset of Artificial Intelligence (AI) that involves the development of algorithms and statistical models that enable computers to learn from and make predictions or decisions without being explicitly programmed.

2. Types of ML algorithms?

There are several types of ML algorithms, including supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning.

- **Supervised Learning**: Algorithms that learn from labeled data and make predictions on new, unseen data. Examples include linear regression, logistic regression, decision trees, and support vector machines.
- **Unsupervised Learning**: Algorithms that learn from unlabeled data and find patterns or structure in the data. Examples include k-means, hierarchical clustering, and principal component analysis.
- **Semi-Supervised Learning**: Algorithms that combine elements of supervised and unsupervised learning and use a small amount of labeled data with a large amount of unlabeled data.
- **Reinforcement Learning:** Algorithms that learn from interacting with an environment and receiving feedback in the form of rewards or penalties.

3. Supervised and Unsupervised

Supervised learning algorithms learn from labeled data and make predictions on new, unseen data. Unsupervised learning algorithms learn from unlabeled data and find patterns or structure in the data.

s

4. Regression VS Classification

Regression algorithms predict continuous values, such as the price of a house. Classification algorithms predict discrete values, such as whether an email is spam or not.

5. Clustering vs association rule mining

Clustering is a technique for grouping similar objects together. Association rule mining is a technique for finding relationships between variables in a dataset. Clustering is an unsupervised learning technique, while association rule mining is a supervised learning technique.

6. Cost function vs Loss function

A cost function is a function that measures the difference between the predicted and actual values of a model. A loss function is a specific type of cost function used in optimization problems.

7. R square vs MSE vs MAE

- **R-square:** A statistical measure of the goodness of fit of a model, it ranges from 0 to 1, 1 being a perfect fit.
- **MSE (Mean Squared Error):** A measure of the average squared difference between the predicted and actual values of a model.
- **MAE (Mean Absolute Error**): A measure of the average absolute difference between the predicted and actual values of a model.

8. What is P value?

A P-value is a statistical measure that represents the probability that the results of a study occurred by chance. A P-value of less than 0.05 is commonly used as a threshold for indicating statistical significance.

9. what does cost function says?

- A cost function is a mathematical function that measures the difference or "cost" between the predicted values of a model and the actual values of the target variable. In machine learning, cost functions are used to evaluate the performance of a model and to guide the optimization process for finding the best set of model parameters.

- For example, in linear regression, the most common cost function is the mean squared error (MSE), which measures the average squared difference between the predicted and actual values of the target variable. The goal of training a linear regression model is to find the set of model parameters that minimize the MSE.

- Similarly, in logistic regression, the most common cost function is the cross-entropy loss, which measures the difference between the predicted and actual probability of the target variable being a certain class.

- In short, cost function is a way to evaluate the performance of a model, it tells us how well our model is doing. If cost function is high that means our model is not doing well and we need to optimize it further.

10.   R square vs Cost function

- R-squared (R2) and cost function are two different measures that are used to evaluate the performance of a model, although they are related.

- R-squared is a statistical measure of the goodness of fit of a model, it ranges from 0 to 1, with 1 being a perfect fit. It measures the proportion of the variance in the target variable that is explained by the model.

- A cost function, on the other hand, measures the difference or "cost" between the predicted values of a model and the actual values of the target variable. It is a mathematical function that is used to guide the optimization process for finding the best set of model parameters.

- So, while R-squared tells us how well our model is fitting the data, cost function tells us how well our model is doing in terms of predicting the target variable.

- It is also possible to use R-squared to compare different models, when the models are trying to fit the same data, R-squared can help us to understand which model is fitting the data better. While cost function is used to evaluate the performance of a single model.