1. What is ML (Machine Learning)?

Machine Learning is a subset of AI that enables systems to learn from data and improve their performance without being explicitly programmed.

2. Types of ML

- **Supervised Learning**: Learns from labeled data (e.g., classification, regression).

- **Unsupervised Learning**: Learns from unlabeled data (e.g., clustering, dimensionality reduction).

- **Reinforcement Learning**: Learns by interacting with the environment and receiving feedback in the form of rewards or penalties.

3. Difference Between Classification and Regression

- **Classification**: Predicts discrete output labels (e.g., spam or not spam).

- **Regression**: Predicts continuous values (e.g., house price).

4. What is a Hypothesis (Hypothesis Class)?

A hypothesis is a possible function or model that maps inputs to outputs.
The hypothesis class is the set of all possible models a learning algorithm can choose from.

5. What is a Reinforcement Learning Problem?

It's a learning problem where an agent learns to make decisions by performing actions and receiving rewards or penalties from the environment.

6. What is Supervised Learning?

A learning approach where the model is trained on labeled data, meaning input-output pairs are known.

7. What is VC Dimension?

VC (Vapnik-Chervonenkis) dimension is a measure of the capacity or complexity of a model — specifically, the largest number of points that can be shattered (correctly classified in all possible ways) by the model.

8. Explain Model Selection

Model selection involves choosing the best model from a set of candidates based on performance metrics like accuracy, precision, recall, etc.

9. What is Regularization?

Regularization is a technique to prevent overfitting by adding a penalty term to the loss function. Common types: L1 (Lasso), L2 (Ridge).

10. What is PAC (Probably Approximately Correct) Learning?

PAC learning is a theoretical framework that defines how a learning algorithm can learn a good approximation of the target function with high probability using a limited amount of training data.