1. What does one mean by the term "machine learning"?

**Machine learning**: A field of AI where systems learn patterns from data to make predictions or decisions without being explicitly programmed.

2.Can you think of 4 distinct types of issues where it shines?

**Types of issues**:

* Image recognition
* Speech recognition
* Fraud detection
* Recommendation systems

3.What is a labeled training set, and how does it work?

**Labeled training set**: A dataset where each input has a known output; the model learns to map inputs to outputs.

4.What are the two most important tasks that are supervised?

**Supervised tasks**:

* Classification
* Regression

5.Can you think of four examples of unsupervised tasks?

**Unsupervised tasks**:

* Clustering
* Dimensionality reduction
* Anomaly detection
* Association rule learning

6.State the machine learning model that would be best to make a robot walk through various unfamiliar terrains?

**Robot walking model**: Reinforcement learning model.

7.Which algorithm will you use to divide your customers into different groups?

**Dividing customers**: Clustering algorithm (e.g., K-Means).

8.Will you consider the problem of spam detection to be a supervised or unsupervised learning problem?

**Spam detection**: Supervised learning problem.

9.What is the concept of an online learning system?

**Online learning system**: A model that updates continuously as new data arrives.

10.What is out-of-core learning, and how does it differ from core learning?

**Out-of-core learning**: Learning from data that cannot fit in memory, unlike core learning which uses in-memory datasets.

11.What kind of learning algorithm makes predictions using a similarity measure?

**Similarity-based prediction**: Instance-based learning (e.g., K-Nearest Neighbors).

12.What's the difference between a model parameter and a hyperparameter in a learning algorithm?

**Model parameter vs hyperparameter**:

* Parameter: Learned from data (e.g., weights in linear regression)
* Hyperparameter: Set before training (e.g., learning rate, number of trees)

13.What are the criteria that model-based learning algorithms look for? What is the most popular method they use to achieve success? What method do they use to make predictions?

**Model-based learning criteria**: Find the model that best fits data; use **optimization** to learn parameters; make predictions using the trained model.

14.Can you name four of the most important Machine Learning challenges?

**ML challenges**:

* Overfitting
* Underfitting
* Data quality and quantity
* Feature selection

15.What happens if the model performs well on the training data but fails to generalize the results to new situations? Can you think of three different options?

**Model performs well on training but fails to generalize**: Overfitting; options:

* Regularization
* More training data
* Simpler model

16.What exactly is a test set, and why would you need one?

**Test set**: A separate dataset used to evaluate final model performance on unseen data.

17.What is a validation set's purpose?

**Validation set purpose**: Tune hyperparameters and evaluate intermediate model performance.

18.What precisely is the train-dev kit, when will you need it, how do you put it to use?

**Train-dev kit**: A split of training data used for model development and debugging before testing; helps prevent overfitting to the test set.

19.What could go wrong if you use the test set to tune hyperparameters?

**Using test set for hyperparameter tuning**: Leads to data leakage and overly optimistic performance estimates.