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

A: A machine tries to fit the given data into a model. This phenomenon is called machine learning.

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

A: User recommendations, fraud detection, speech recognition, stock price forecast

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

A: A data set containing both input and output, and can be used to train a model is called a labeled training set. When a labeled training set is used, the machine tries to identify the relation between input and output.

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

A: Regression and classification

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

A: Clustering, dimensionality reduction, visualization and association rule learning.

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

A: Reinforced learning

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

A: Clustering

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

A: Supervised learning

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

A: When a machine is provided with data from online sources continuously, it is called online learning system

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

A: Out-of-core learning is a type of learning used when the data does not fit in the RAM.

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

A: Instance based learning

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

A: Hyperparameter relates to all instances of a learning algorithm while model parameter relates to only one instance.

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?

A: They look for the optimal parameters. Most of the time, a cost function or a similar function is minimised to achieve this. These optimal parameters along with the models are used to make predictions.

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

A: Overfitting, Underfitting, Poor quality of the data and less amount of training data

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?

A: This is called overfitting of the data. This might happen due to the existence of outliers, or bad sampling of data for training or a complex ML algorithm.

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

A: A test set is used to check the generalizability of a model.

17.What is a validation set's purpose?

A: Generally it is used to model tuning. Use of a validation set helps in avoiding training the model using a test set accidentally.

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

A: While training a model, data is usually split into 3 data sets. A training set is used to train models while dev set is used to compare these models

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

A: The purpose of a test set is to verify if a trained model works on new data or not. This purpose might be compromised if we use the test set for model tuning as we are indirectly training the model for the test set.