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

Solution - Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.

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

Solution – Machine learning is used in internet search engines, email filters to sort out spam, websites to make personalised recommendations, banking software to detect unusual transactions, and lots of apps on our phones such as voice recognition.

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

Solution - The training set is used to train the algorithm, and then you use the trained model on the test set to predict the response variable values that are already known. A labelled training set refers to a dataset having a label output along with its features.

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

Solution – The two most common supervised learning tasks are regression and classification. In a regression problem we our prediciton is a scalar value. When we're trying to solve a classification problem, our output is either 1 or 0.

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

Solution - Four common unsupervised tasks inclused clustering, visualization, dimensionality reduction, and association rule learning.

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

Solution – The best Machine Learning algorithm to allow a robot to walk in unknown terrain is Reinforced Learning, where the robot can learn from response of the terrain to optimize itself.

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

Solution – Clustering algorithm

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

Solution - Spam detection is a supervised machine learning problem.

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

Solution – Online learning refers to instruction that is delivered electronically through various multimedia and Internet platforms and applications.

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

Solution – Out-of-core learning refers to the machine learning algorithms working with data that cannot fit into a single machine's memory but can easily fit into some data storage, such as a local hard disk or web repository.

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

Solution – Learning algorithm that relies on a similarity measure to make predictions is instance-based algorithm.

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

Solution – Model Parameters: These are the parameters in the model that must be determined using the training data set. These are the fitted parameters. Hyperparameters: These are adjustable parameters that must be tuned in-order to obtain a model with optimal performance.

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?

Solution- Model based learning algorithm search for the optimal value of parameters in a model that will give the best results for the new instances. We often use a cost function or like determine what the parameter value has to be in order to minimize the function. The model makes prediction by using the value of the new instance and the parameters in its function.

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

Solution- Four main challenges in Machine Learning include overfitting the data (using a model too complicated), underfitting the data (using a simple model), lacking in data and nonrepresentative 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?

Solution- If the model performs poorly to new instances, then it has overfit on the training data. To solve this, we can do any of the following three: get more data, implement a simpler model, or eliminate outliers or noise from the existing data set.

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

Solution- Test set is a set that you test your model (fit using training data) to see how it performs. Test set is necessary so that you can determine how good (or bad) your model performs.

17.What is a validation set's purpose?

Solution- Validation set is a set used to compare between different training models

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

Solution- Training process emits the parameters of a model and hence the sole purpose of training data is to decide about which parameters to pick given huge options to choose from.

The goal of dev-set is to rank the models in term of their accuracy and helps us decide which model to proceed further with.

In the process of building a machine learning project, we made a lot of decisions from parameters of the model to the different choice of models. Choices of parameters are made by learning algorithm and choice of model is made by us by intentionally picking the best model performing well on dev-set. So when you have successfully consumed train and dev set while building a machine learning system you are left with the final best model out of all the ideas you tried.

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

Solution- If you tune hyperparameters using the test sets, then it may not perform well on the out-of-sample data because the model is tuned just for that specific set.