1. What does one mean by the term &quot;machine learning&quot;?

A: Machine learning is a branch of [artificial intelligence (AI)](https://www.ibm.com/in-en/topics/artificial-intelligence) 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?

### Speech & Image Recognition

### Chatbot

### Google Translation

### Self-driving cars

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

A: In machine learning, data labeling is the process of identifying raw data (images, text files, videos, etc.) and adding one or more meaningful and informative labels to provide context so that a machine learning model can learn from it

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, 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?

A: Reinforced Learning

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

A: Clustering algorithm

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

problem?

A: supervised machine learning problem

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

A:

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

A: It is a way to train your model on data that cannot fit your core memory.” 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?

A: instance-based algorithm

12.What&#39;s the difference between a model parameter and a hyperparameter in a learning

algorithm?

A: Basically, parameters are the ones that the “model” uses to make predictions etc. For example, the weight coefficients in a linear regression model. Hyperparameters are the ones that help with the learning process. For example, number of clusters in K-Means, shrinkage factor in Ridge Regression.

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:

* the optimal value of parameters in a model that will give the best results for the new instances.
* Decision tree

A decision tree is **a non-parametric supervised** learning algorithm, which is utilized for both classification and regression tasks. It has a hierarchical, tree structure, which consists of a root node, branches, internal nodes, and leaf nodes.

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

A:

* overfitting the data (using a model too complicated)
* underfitting the data (using a simple model)
* lacking in data
* nonrepresentative data.

15.What happens if the model performs well on the training data but fails to generalize the results

to new situations?

Overfitting

Can you think of three different options?

* A student does well mugging up all the questions , but may face challenges in application based questions.
* Training data set based on price of houses of a particular location may differ with respect to location.
* Data set of food habit of a particular age gr may vary with other age gr.

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

A: A test set is a portion of a data set used in data mining to assess the likely future performance of a single prediction or classification model that has been selected from among competing models, based on its performance with the validation set.

17.What is a validation set&#39;s purpose?

A: A validation set is a set of data used to train artificial intelligence (AI) with the goal of finding and optimizing the best model to solve a given problem. Validation sets are also known as dev sets.

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

A: 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. Using Dev set we rank all our models in terms of their accuracy and pick the best performing model.

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

* lack of data,
* poor data quality,
* nonrepresentative data
* uninformative features
* excessively simple models that underfit the training data.
* complex models that overfit the data.