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

**Ans :-** Machine Learning is basically a subset of artificial intelligence which provides us a powerful tools to visualize and analyse the data. By it we can also make some prediction and forecasting against the new dataset.

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

**Ans:-** Four types of issues where ML shines are:

* Problems on spam detection in emails.
* Problems on Credit card fraud transaction’s.
* Problems on Cancer Diagnosis prediction .
* Problems on automatically driving cars.

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

**Ans:-** During the model training we need to split the dataset into two parts i.e. Training data and test data **.** 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 .After training the model on test data we need to predict the efficiency of the model.

1. **What are the two most important tasks that are supervised?**

**Ans:-** The two most important supervised tasks are **regression** and **classification.**

1. **Can you think of four examples of unsupervised tasks?**

**Ans:-** Four common unsupervised tasks includes clustering, visualization, dimensionality reduction , and association rule learning.

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

**Ans:-** Reinforcement learning is a system where an "agent" observes the environment, selects and performs actions, then receives a reward or punishment based on the result of the action.

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

**Ans:-** We can use classification algorithm approach to divide the customers in different groups.

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

**Ans:-** The problem of spam detection to be a supervised learning problem because in this we know that the labels are known(Spam and ham messages).

1. **What is the concept of an online learning system?**

**Ans:-** Online learning is an approach that ingests data one observation at a time.

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

**Ans:-** Out-of-core learning is used when a dataset is too large to fit into a computer's memory. The algorithm loads part of the data, runs a training step, then repeats the process until it has run on all the data.

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

**Ans:-** Instance-based learning algorithms use a measure of similarity to generalize to new cases.

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

**Ans:-** Model parameters are configuration variables that are internal to the model, and a model learns them on its own , Parameters are essential for making predictions. and Hyper parameters are those parameters that are explicitly defined by the user to control the learning process. Hyper parameters are essential for optimizing the model.

1. **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?**

**Ans:-** 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 similar to determine what the parameter value has to be in order to minimize the function. This involves using algorithms like linear regression, logistic regression, and random forest.

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

**Ans:- Four of the most important Machine Learning challenges are:-**

* Your dataset is full of missing values, outliers, and noise (poor measurements).
* The features in your dataset are irrelevant. Garbage in, garbage out.
* Overfitting happens
* Under fitting happens

1. **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?**

**Ans:-** This is a case where the model is overfitting the training data.

**Three different options are:-**

* Select a more powerful model with more parameters.
* We could gather more data.
* Use feature engineering to feed better features to the model.
* We can reduce noisiness in the data by fixing errors and removing outliers.

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

**Ans:-** During the model training we need to split the dataset into two parts i.e. Training data and test data **.** Our model is training with the training set. Then we use the model to run predictions on the test set.

1. **What is a validation set's purpose?**

**Ans:-** The validation set is used to evaluate the models in order to perform model selection. The validation set is also used for hyper parameter tuning.

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

**Ans:-** The goal of dev-set is to rank the models in term of their accuracy and it helps us to decide which model to proceed further with. First training algorithm makes the choice for optimal parameters and then those parameters are used on dev data to help us find best model architecture.

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

**Ans:-** The model has chance to "see" the test data and to develop a bias towards this test data. You actually lose the possibility to find out how good your model would actually be on unseen data