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

- Machine learning stands for a process where machine tries to find out the relationship between variables or features which includes the learning parameter in it. The output is combination of prediction which is given to machine as input in the form of data. It is a way for computer programs to improve their performance on a task over time given more data.

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

- 1. In the Automation, where the work which is done manually it can be automated with the help of machine learning.

2. In the field data prediction such as Stocks, Weather etc.

3. spam detection in email

4. cancer diagnosis

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

- Labeled training set is a type of data where input ie features are in the format of label. For Example if we consider house price prediction use case then in this use case the input features such as TV, Radio are labelled ones. These do contain values in them.

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

- The two most common supervised learning tasks are regression and classification. In a regression problem , 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?**

Ex 1 :- Semantic Clustering

Ex 2 :- Recommendation system

Ex 3 :- Market based analysis

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

- Reinforcement learning - 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

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

- Clustering based algorithm like k- means clustering.

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

- Supervised learning because we do know the output and input features whether it is spam or ham.

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

- An online learning system learns from new data on-the-fly. As a result, the system is trained incrementally either by using one example at a time or using a mini-batch approach. This keeps each learning step cheap and memory efficient.

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

- 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.

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

- Instance-based learning algorithms use a measure of similarity to generalize to new cases. In an instance-based learning system, the algorithm learns the examples by heart, then uses the similarity measure to generalize

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

- Model parameters help us to change or predict the values on the basis of trained model and hyperparameter helps to retrain the model by tuning few of its parameters to achieve better accuracy.

**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?**

- The goal for a model-based algorithm is to be able to generalize to new examples. To do this, model based algorithms search for optimal values for the model's parameters, often called theta. This searching, or "learning", is what machine learning is all about. Model-based system learn by minimizing a cost function that measures how bad the system is at making predicitons on new data, plus a penalty for model complexity if the model is regularized. To make a prediction, a new instance's features are fed into a hypothesis function which uses the minimized theta found by repeatedly running the cost function.

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

- 1. Images, videos create a challenge for a model to train.

2. Huge number of data challenge in computation.

3. Data underfitting or overfitting happens while training.

4. Accuracy or Prediction becomes tough for huge data when its been trained.

**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?**

- It is the case of overfitting, where on the training data the prediction is good but on test data the prediction fails or is of poor one. Here , we can try with

1. Cross validation

2. Regularization

3. Data Exploration or augmentation.

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

- Test data is required when the data given to the model is been trained and after training the trained data of the model should be checked with different input as how the model is behaving.

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

- Validation data is required to validate the trained model which is been trained with the trained data.

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

- train-dev kit helps us to identify better model. It ranks the models in term of their accuracy and helps us decide which model to proceed further with. We can implement this while building project which includes a particular file example named as model factory.

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

- There will be no difference in the output because we use hyperparameters on the training set to retrain as well as tune our model and then get a best possible model out of it. But when this happens with test data output doesn’t differ and we actually lose the possibility to find out how good your model would actually be on unseen data.