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

ANSWER- Machine Learning is the science of programming computers so they can learn from data. Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed.

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

ANSWER- Four distinctive type of issues where it shines:

1. Problems for which existing solutions require a lot of hand-tuning or long lists of rules.

2. Complex problems for which there is no good solution at all using a traditional approach.

3. Fluctuating environments, a Machine Learning system can adapt to new data.

4. Getting insights about complex problems and large amounts of data.

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

ANSWER- In a labeled training ,there are few lables attached to each data . For exapmle , spam filter in e-mails , the spam mails have their own labels so new lable e-mail coming will be checked with present mail labels and then accordingly putted into spam filter.

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

ANSWER- Two most important tasks that are supervised:

1. Classification :

    The model is trained with many examples to classify data ,it may uses lables and machine learning algorithm like Logistic Regression.

2. Predict a Target :

    Predicting a taget numeric value is also a crucial task in machine learning . For example, Price of a car with a given set of Features.

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

ANSWER- Examples are:

1. Hierarchical Cluster Analysis (HCA)

2. Principal Component Analysis (PCA)

3. t-distributed Stochastic Neighbor Embedding (t-SNE)

4. Isolation Forest

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

ANSWER- Reinforcement Learning would be best to make a robot walk through various unfamiliar terrains as in this system first observes the enviornment and takes action according to policy and then get reward and penality . If penality then changes the action in policy as well.

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

ANSWER- Clustering Algorithm , a type of Unsupervised algorithm divides customers into different groups on the basis of their feature.

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

ANSWER- Spam detection would be considered as supervised learning problem .

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

ANSWER- In online learning, the system is trained incrementally by feeding it data instances sequentially, either individually or by small groups called mini-batches. It is great for systems that receive data as a continuous flow and need to adapt to change rapidly or autonomously.

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

ANSWER- Online learning algorithms can also be used to train systems on huge datasets that cannot fit in one machine's main memory ,this is called out-of-core learning. Out-of-core learning is usually done offline. The algorithm loads part of the data, runs a training step on that data, and repeats the process until it has run on all of the data. Core Learning enables Developers to integrate their Machine learning models into iOS applications whereas Out of core learning refers to the machine Learning algorithm working with data that cannot fit into a single machine's memory but can easily fit into some data storage.

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

ANSWER- Instance – Based Machine Learning Algorithm makes prediction using a similarity measure.

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

ANSWER- **Model parameter** are the parameters in the model that must be determined using training data set. These are the fitted parameters . Whereas The **Hyper Parameter** 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?

ANSWER- Model – based learning algorithm look for **Theta,** optimal values for the model's parameter.

Minimization of loss and accuracy or regularization or we can say minimizing the difference the prediction of the model and the actual values of the output variable.

By using the learned model to estimate the output variable for new variable feature , prediction are done.

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

ANSWER- The four most important challenges of Machine Learning are :

1. Insufficient Quantity Of Training Data -

    For a machine learning algorithm to work properly you need a lot of data . It need a thousands of examples even for a simple problem, for complex problem like speech recoginition and image recoginition it requires a lot of data to make machine learning algorithm work properly.

2. Nonrepresentative Training Data -

    By using a nonrepresentative training set, the trained model is unlikely to make accurate predictions.

3. Poor-Quality Data -

    If training data is full of errors, outliers, and noise ,it will make it harder for the system to detect the underlying patterns, so your system is less likely to perform well. It is often well worth the effort to spend time cleaning up your training data.

4. Irrelevant Features -

    System will only be capable of learning if the training data contains enough relevant features and not too many irrelevant ones. A critical part of the success of a Machine Learning project is coming up with a good set of features to train on.

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?

ANSWER- Three different situations are :

1. Poor Performance is one of the situations , in which when model algorithm applied to training data it gives high accuracywhereas in new situation it fails.

2. Less control over noise and outliers, the outliers in new data are not properly managed by the model .

3. The model might not be able to handle the change in the input data , it is not able to adapt quickly with respect to new data.

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

ANSWER- The given dataset is spilit into train and test dataset . The train dataset acquires 80% of total dataset whereas 20% is test dataset. When machine Learning model is created then all prediction is done on train dataset while training the model. While on training phase of model train dataset is used. To test the accuracy of model prediction the test dataset is used . Both train and test dataset are randomly generated . Basically train dataset checks the accuracy of modle prediction.

17.What is a validation set's purpose?

ANSWER- Validation set evaluates the performance of a model during the training process and assist in the selection of hyperparameters and model architecture. It helps to know that the model can perform well real- world sceenarios as the validation is totally different from training set and representative of the real – world data distribution. It is often used for early stopping techniques.

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

ANSWER- train- dev kit , Training development dataset , the dataset that is used for both training and initial model development. It can be used when the available data is limited, and a seperate validation set is not feasible due to data scaricity.

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

ANSWER- if we use the test set to tune hyperparameters then there might be chances of overfitting to test set , there might be Lack of Generalization, and Inflated performance metrics. To neutralize these issues there must be availability of a valiadation set to tune hyperparameters.