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

Ans. Machine learning is about making systems to learn from data without having to expilicitly program them. The system should learn from experiences(data) and should perform a task given some performance measure.

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

Ans. 1> Ml is good where the existing solution requires lot of fine-tuning(which can take enormous time).

2> complex problems where the traditional system yields no good result

3>fluctuating environment where the system fails to adapt new or continuously changing data.

4> data mining(getting insights from large data becomes easy)

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

Ans. Labeled training set is one in which a desired solution is included for each instance.

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

Ans. Classification and Regression

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

Ans.Clustering,visualization,Dimensionality reduction,anomaly and novelty detection

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

Ans. Reinforcement learning

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

Ans. Depending on whether the group is defined for us or not we can use either classification or clustering respectively.

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

Ans. supervised

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

Ans. Online learning system is the one which can learn incrementally by feeding it data instances sequentially or in small groups.It learns and adapt to changing data.It can be used where data quantity is enormous and we need an automated system that derive solution through this data on its own.

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

Ans. Out of core learning means training the system offline.It is generally done where we are limited by machine’s main memory(owing to huge dataset). The data is feeded in small groups called mini-batches and then incremental learning method is applied to learn from these mini batches.

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

Ans. Instance based learning.

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

Ans. Model parameter will determine what a model will predict given a new instance.for ex for linear regression y=mx+c(m and c are the model parameters whereas hyperparamter is a parameter is a parameter of the learning algorithm itself not of the model(it helps to to control or set the amount of regularization to apply).

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?

Ans. Model based learning looks for an optimal value for the model parameter so that the model can generalize well on new instances.We can use either a cost function to check how bad a model has performed or a fit function for checking how good a model is.For prediction we feed the new instance features into the model’s prediction function using the model paramter found by learning algorithm.

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

Ans. Lack of data,poor quality of data,nonrepresentative data,irrelevant features.

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?

Ans.It is a case of overfitting. We can simplify the model(regularization), get more data, reducing the noise in the training data.

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

Ans. A test set is the one where the trained model is tested to check how well it is generalizing to new instances. It is needed to check the generalization error our model is making before launching the model into production.

17.What is a validation set's purpose?

Ans. It is used to compare models.It makes it possible to select the best model and tune the hyperparameter.

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

Ans. Train-dev set or dev set is a part of training set that is held out. The model is trained on rest of the training set(except dev set) and then the model is evaluated on both the dev set and validation set.it is done to check how well a model is performing or generalizing on new instances if the model is doing good on the training set and not on dev set then the model is likely to overfit.if model performs well on both training and dev set but not in validation set then there is a chances of data mismatch, and we should try to improve the training data to make it look more like the validation and test data.

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

Ans. It can lead to overfitting and the model we launch may perform worse than we expected(the model won’t perform as well on new data).