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

Ans. Machine learning (ML) is a type of artificial intelligence (AI) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so. Machine learning algorithms use historical data as input to predict new output values.

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

Ans. 1. Prediction of sales

1. Classification of flowers
2. House price prediction
3. Weather Prediction

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

Ans. A labeled training set is a collection of data where one of the features of the data indicates the class the training example belongs to. A labeled training set is used in supervised learning algorithms.

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

Ans. Two most common supervised tasks are classification and regression

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

Ans. Common unsupervised tasks include 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?

Ans. Reinforcement learning

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

Ans. Clustering Algorithms

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

Ans. Supervised problem.

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

Ans. 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?

Ans. Out-of-core learning refers to a set of algorithms working with data that cannot fit into the memory of a single computer, but that can easily fit into some data storage such as a local hard disk or web repository.

Core learning involves learning experiences in situations contextually relevant to the particular qualification. Ideally, these graduating majors will indicate proficient achievement of the Philosophy Core learning goals.

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

Ans. Learning algorithm that relies on a similarity measure to make predictions is instance-based algorithm.

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

Ans. Model Parameters: These are the parameters in the model that must be determined using the training data set. These are the fitted parameters.

Hyperparameters: These 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?

Ans. 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 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.

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

* Ans. Not gathering enough data, or sampling noise. Sampling noise means we'll have non-representative data as a result of chance.
* Using a dataset that is not representative of the cases you want to generalize to. This is called sampling bias. For example, if you want to train an algorithm with "cat videos", and all your videos are from YouTube, you're actually training an algorithm to learn about "YouTube cat videos."
* Your dataset is full of missing values, outliers, and noise (poor measurments).
* The features in your dataset are irrelevant. Garbage in, garbage out.
  + Feature selection - choose the most relevant features from your dataset
  + Feature extraction - combine features in your dataset to generate a new, more useful feature
* When your model performs well on the training data, but not on test data, you've over fit your model. Models that suffer from overfitting do not generalize well to new examples. Overfitting happens when the model is too complex relative to the amount and noisiness of the data.
  + Try simplyfying the model by reducing the number of features in the data or constraining the parameters by reducing the degrees of freedom.
  + Gather more training data.
  + Reduce noise in the training data by fixing errors and removing outliers.
* When your model is too simple to learn the underlying structure of the data you've underfit your model.
  + Select a more powerful model with more parameters
  + Use feature engineering to feed better features to the model
  + Reduce the constraints of the model (increase degrees of freedom, reduce regularization parameter, etc.)

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. This is a case where the model is overfitting the training data. To couteract overfitting, we can reduce the complexity of the model by removing features or constraining the parameters. We could gather more data. Finally we can reduce noisiness in the data by fixing errors and removing outliers.

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

Ans. When we want to know how well our model generalizes to new cases we prefer to use a test set instead of actually deploying the system. To build the test set we split the training data (50-50, 60-40, 80-20 are common splits) into a training set and test set. Our model is training with the training set. Then we use the model to run predictions on the test set. Our error rate on the test set is called the generalization error or out-of-sample error. This error tells us how well our model performs on examples it has never seen before.

If the training error is low, but the generalization error is high, it means we're overfitting our model.

17.What is a validation set's purpose?

Ans. Let's say we have a linear model and we want to perform some hyperparameter tuning to reduce the generalization error. One way to do this 100 different models with 100 different hyperparameter values using the training set and finding the generalization error with the test set. You find the best hyperparameter value gives you 5% generalization error.

So you launch the model into production and find you're seeing 15% generalization error. This isn't going as expected. What happened?

The problem is that for each iteration of hyperparameter tuning, you measured the generalization error then updated the model using the same test set. In other words, your produced the best generalization error for the test set. The test set no longer represents cases the model hasn't seen before.

A common solution to this problem is to have a second holdout set called the validation set. You train multiple models with various hyperparameters using the training set, you select the model and hyperparameters that perform best on the validation set, and when you are happy about your model you run a single final test against the test set to get an estimate of the generalization error.

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

Ans. Your model will not be generalizable to new examples.

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

Ans. Cross-validation helps us compare models without wasting too much training data in the validation set.