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

Machine learning is the practice of applying algorithmic models to data, in an iterative manner, so

that your computer discovers hidden patterns or trends that you can use to make predictions. It’s

also called algorithmic learning.

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

Spam filter, speech recognition, Recommendation engines, Retention modelling, Predictive maintenance, medical diagnoses, Fraud detection

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

In supervised learning, the training data you feed to the algorithm includes the desired solutions or output values, called labels. To train any system, you need to give the system many examples, including both their features or predictors and their target or labels. These algorithms learn from known features of that data to produce an output model that successfully predicts labels for new incoming, unlabeled data points.

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

**Classification** - Ex The spam filter is a good example of this: it is trained with many example emails along with their class (spam or ham), and it must learn how to classify new emails.

**Regression** – Ex predict a target numeric value- the price of a car, given a set of features (mileage, age, brand, etc.) called predictors.

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

1. **clustering algorithm** to try to detect groups of similar visitors to your blog.

2**. visualization algorithms**- you feed them a lot of complex and unlabeled data, and they output a 2D or 3D representation of your data that can easily be plotted.

3.**dimensionality reduction**, in which the goal is to simplify the data without losing too much information. One way to do this is to merge several correlated features into one.

4. **anomaly detection** — for example, detecting unusual credit

card transactions to prevent fraud, from a dataset before feeding it to another learning algorithm. The system is trained with normal instances, and when it sees a new instance, it can tell whether it looks like a normal one or whether it is likely an anomaly.

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

**Reinforcement Learning** -The learning system, called an agent in this context, can observe the environment, select, and perform actions, and get rewards in return (or penalties in the form of negative rewards. It must then learn by itself what is the best strategy, called a policy, to get the most reward over time. A policy defines what action the agent should choose when it is in each situation.

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

**Clustering algorithm** to detect groups of similar customers. If the categories of customers is known then I would prefer classification algorithm.

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

I would consider the spam detection as a supervised learning problem.

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

In online learning, you train the system incrementally by feeding it data instances sequentially, either

individually or by small groups called mini batches. 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 algorithms can handle vast quantities of data that cannot fit in a computer’s main

memory. 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.

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

An instance-based learning system learns the training data by heart; then, when given a new instance, it uses a similarity measure to find the most similar learned instances and uses them to make predictions.

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

**Model parameters are configuration variables that are internal to the model, and a model learns them on its own.** A model has one or more model parameters that determine what it will predict given a new instance.(e.g., the slope of a line and intercept of the line). A learning algorithm tries to find optimal values for these parameters such that the model generalizes well to new instances.

**Hyperparameters are those parameters that are explicitly defined by the user to control the learning process.** A hyperparameter is a parameter of the learning algorithm itself, not of the model (e.g., learning rate, number of iterations, number of hidden layers, number of hidden units, choice of activation functions). These are parameters that control the ultimate parameters.

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?

Model-based learning algorithms search for an optimal value for the model parameters such that the model will generalize well to new instances.

We usually train such systems by minimizing the cost function that measures how bad the system is at making predictions on the training data, plus a penalty for model complexity if the model is regularized.

To make predictions, we feed the new instance’s features into the model’s prediction function, using the parameter values found by the learning algorithm.

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

The two things that can go wrong are “bad algorithm” and “bad data.” The system will not perform well if your training set is too small, or if the data is not representative, noisy, or polluted with irrelevant features (garbage in, garbage out). Lastly, your model needs to be neither too simple (in which case it will underfit) nor too complex (in which case it will overfit).

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?

In Machine Learning this is called ***overfitting***: it means that the model performs well on the training data, but it does not generalize well. The three different solutions are

1. simplify the model by selecting one with fewer parameters (e.g., a linear model rather than a high-degree polynomial model), by reducing the number of attributes in the training data or by constraining the model.
2. gather more training data
3. reduce the noise in the training data (e.g., fix data errors and remove outliers)

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

A test set is used to estimate the generalization error that a model will make on new instances, before the model is launched in production. We split the training data into a training set and test set. Our model is trained with a 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?

A validation set is used to compare models. It makes it possible to select the best model and tune the hyperparameters.  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?

The goal of **dev-set** is to rank the models in terms of their accuracy and helps us decide which model to proceed further with. If the training set and dev sets have different distributions, it is good practice to introduce a train-dev set that has the same distribution as the training set. This train-dev set will be used to measure how much the model is overfitting.

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

If you tune hyperparameters using the test set, you risk overfitting the test set, and the generalization error you measure will be optimistic (you may launch a model that performs worse than you expect).