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

Machine learning Popularly known as ML is a branch 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?

* **Image Recognition**: Image recognition is one of the most common applications of machine learning. It is used to identify objects, persons, places, digital images, etc. The popular use case of image recognition and face detection is **Automatic friend tagging suggestion**.
* **Speech Recognition**: While using Google, we get an option of **Search by voice**, it comes under speech recognition, and it's a popular application of Machine Learning. Speech recognition is a process of converting voice instructions into text, and it is also known as **Speech to text**, or **Computer based speech recognition** At present, machine learning algorithms are widely used by various applications of speech recognition. Google assistant, Siri, Cortana, and Alexa are using speech recognition technology to follow the voice instructions.
* **Traffic prediction**: It predicts the traffic conditions such as whether traffic is cleared, slow-moving, or heavily congested with the help of two ways: Real Time location of the vehicle form Google Map app and sensors Average time has taken on past days at the same time.
* **Product recommendations**: Machine learning is widely used by various e-commerce and entertainment companies such as Amazon, Netflix, etc., for product recommendation to the user.
* **Credit card fraud detection**
* **Autonomous vehiccles**

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

The training set is **used to train the algorithm, and then you use the trained model on the test set to predict the response variable values that are already known**. The final step is to compare the predicted responses against the actual (observed) responses to see how close they are

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

Regression & Classification

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

**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?

Reinforcement learning

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

Clustering

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

Supervised

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

Online learning system is a learning system in which the machine learns continously, as data is given in small streams continuously.

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

 Out-of-core learning system is a system that can handle data that cannot fit into your computer memory. It uses online learning system to feed data in small bits.

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

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?

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?

 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. The model makes prediction by using the value of the new instance and the parameters in its function.

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

1. Overfitting the Data (using a model too complicated)
2. Underfitting the data (using a simple model)
3. Lacking in Data
4. Non Representative Data

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?

* Get more data
* Implement a simpler model
* Eliminate outliers or noise from the existing data set.
* Regularization

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

Test set is a set to test your model (fit using training data) to see how it performs.Test set is necessary to determine how good (or bad) a model performs.

17.What is a validation set's purpose?

Validation set is a set used to compare between different training models.

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 term of their accuracy and helps us decide which model to proceed further with. Using Dev set we rank all our models in terms of their accuracy and pick the best performing model. i.e. dev set ranks models similar to a search engine like google rank pages and then pick the top model and hence act as a filter to remove bad models.

t is **the set of data that is used to train and make the model learn the hidden features/patterns in the data**

A validation data set is a data-set of examples used **to tune the hyperparameters (i.e. the architecture) of a classifier**. It is sometimes also called the development set or the "dev set". An example of a hyperparameter for artificial neural networks includes the number of hidden units in each layer.

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

If you tune hyperparameters using the test sets, then it may not perform well on the out-of-sample data because the model is tuned just for that specific set.