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

Ans: 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?

Ans: The following are some of the issues where Machine Learning can be used:

- **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.

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

Ans: We split up the data containing known response variable values into two pieces. 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.

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

Ans: The two most common supervised learning tasks are **Regression** and **Classification**.

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

Ans: Four common Unsupervised Tasks included **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: The best Machine Learning algorithm to allow a Robot to walk in unfamiliar terrains is **Reinforced Learning**, where the robot can learn from response of the terrain to optimize itself.

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

Ans: The Best Algorithm to Segment Customers into different groups is either **Supervised Learning** (if the groups have known labels) or **Unsupervised Learning** (if there are no group labels).

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

Ans: Spam detection is a Supervised Machine Learning problem because the labels are known (spam or no spam).

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

Ans: 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?

Ans: 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?

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 parameter determines how a model will predict given a new instance. Model usually has more than one parameter (i.e. slope of a linear model). Hyperparameter is a parameter for the learning algorithm, not of a model.

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 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?

Ans: Four main challenges in Machine Learning include the following:

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?

Ans: If the model performs poorly to new instances, then it has overfitted on the training data. To solve this, we can do any of the following three:

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

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

Ans: 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?

Ans: 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?

Ans: Cross-validation is a tool to compare models without needing a separate validation set. It is preferred over validation set because we can save from breaking of part of the training set to create a validation set, as having more data is valuable regardless.

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

Ans: 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.