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

Machine Learning is the field of study that gives computers the capability to learn without being explicitly programmed

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

1. 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

2. Speech Recognition

While using Google, we get an option of "Search by voice," it comes under speech recognition, and it' shortest route and predicts the traffic conditions.

3. Traffic prediction:

If we want to visit a new place, we take help of Google Maps, which shows us the correct path with the shortest route and predicts the traffic conditions. s a popular application of machine learning.

4. 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?

The separation of the data into a training portion and a test portion is the way the algorithm learns. You 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. The final step is to compare the predicted responses against the actual (observed) responses to see how close they are. The difference is the test error metric. Depending on the test error, you can go back to refine the model and repeat the process until you’re satisfied with the accuracy.

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

The two most common supervised learning tasks are regression and classification

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

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?

The best Machine Learning algorithm to allow a robot to walk in unknown terrain 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?

k-means clustering algorithm

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

Spam detection is a supervised machine learning problem

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

Online machine learning is a method of machine learning in which data becomes available in a sequential order and is used to update the best predictor for future data at each step, as opposed to batch learning techniques which generate the best predictor by learning on the entire training data set.

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

The term out-of-core typically refers to processing data that is too large to fit into a computer’s main memory

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 are estimated based on the data during model training and model hyperparameters are set manually and are used in processes to help estimate model 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?

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 .

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

* Poor Quality of Data
* Underfitting of Training Data
* Overfitting of Training Data
* Lack of Training 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?

Underfitting refers to a model that can neither model the training data nor generalize to new data. An underfit machine learning model is not a suitable model and will be obvious as it will have poor performance on the training data.

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

The test set is a separate set of data used to test the model after completing the training.

It provides an unbiased final model performance metric in terms of accuracy, precision, etc. To put it simply, it answers the question of "How well does the model perform?"

17.What is a validation set's purpose?

The validation set is a set of data, separate from the training set, that is used to validate our model performance during training

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

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

* poor data quality
* nonrepresentative data
* uninformative features
* excessively simple models that underfit the training data
* complex models that overfit the data