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

Machine Learning is the branch of Artificial Intelligence that gives the machines the ability to automatically learn from data and predict outcomes more accurately with minimum human intervention.

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

Machine learning is great for problems whose solution requires a great deal of work or along list of rules, complex problems that are hard to get a solution using traditional method, fluctuating environments and getting insights about complex problems and large data.

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

Labeled is used in Supervised Learning algorithm. Labeled training set is used to train the model. This trained model is used on the test set to predict the response variable which is already known.

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

The two most important supervised tasks are Regression and classification.

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

Four examples of unsupervised tasks are 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 model to make a robot walk through various unfamiliar terrains is Reinforcement Learning.

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

K-means clustering is the best algorithm for customer segmentation. Clustering algorithm is the process of classifying the customers into the same segment.

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

Spam detection is a supervised learning which predicts Spam or not a Spam.

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. It is used to update the best predictor for future data at each step.

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

out-of-core learning is a machine learning algorithm working with data that cannot fit into a single machine’s memory, but it can fit in single data storage(eg.Local hard disk, web repository) . It uses online learning system to feed in the data in small bits.

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

Instance based algorithms is a kind of algorithm that makes predictions using similarity measure.

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

Model Parameters are internal to the model and estimated from the data automatically. These are fitted parameters

Hyperparameters are set manually and are used in optimization of the model and helps in estimating the model parameters. These are adjustable 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 optimal values of the parameters in the model that gives best results for the new instances.

Cost Function is used to determine what the parameter value has to be to minimise the function.

The model makes predictions using the new values of new instances and parameters in its functions.

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

Overfitting and Under fitting, Poor Quality of Data, Lack of Training Data, and Customer Segmentation are the most important Machine Learning challenges.

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?

Overfit on the training data is the cause of the poor performance in new instances.

To solve this problem: Add more training data, implement simple models, eliminate noise, outliers from the existing data set.

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

Test set is used to test your machine learning model to see how it performs. We need test set to see how good or bad the model performs.

17.What is a validation set's purpose?

validation set is a set of data separate from training set which is used to validate the model performance during training. Validation sets are also known as dev sets. Validation set is used for hyperparameter tuning.

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

Train-Dev kit data is the random samples taken from the Training set. Train-dev kit comes from the same distribution as training set but not used for training.

When the training and dev- test set doesn’t come from the same distribution.

Split some random samples of data from training and name it as Train-Dev set.

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

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