**How would you define Machine Learning?**

The Machine learning is learning patterns by machine itself and find hidden relation between attributes. It doesn’t require to manually write rules to extract particular pattern from any problem.

**Can you name four types of problems where it shines?**

Facial recognition, bitcoin prediction, household price estimations, spam email detection.

**What is a labeled training set?**

Label trainset is used for fed to supervised learning where each instance is label. In case of spam filter every email is marked spam or ham manually before fed it to algorithm.

**What are the two most common supervised tasks?**

KNN, Linear regression.

**Can you name four common unsupervised tasks?**

Clustering, Associate learning, principal component analysis, anomaly detection.

**What type of Machine Learning algorithm would you use to allow a robot to walk in various unknown terrains?**

Reinforcement Learning.

**7. What type of algorithm would you use to segment your customers into multiple groups?**

Clustering.

**8. Would you frame the problem of spam detection as a supervised learning problem or an unsupervised learning problem?**

it is supervised learning problem.

**9. What is an online learning system?**

Online learning is term used to learning over to continuously fed data to ML.

**10. What is out-of-core learning?**

Out-of-core learning is when data is huge and cannot be fed entirely at once. This require huge RAM power. This is known as out-of-core learning where small chunks of data are fed instead.

**11. What type of learning algorithm relies on a similarity measure to make predictions?**

Unsupervised learning, clustering, KNN.

**12. What is the difference between a model parameter and a learning algorithm’s hyperparameter?**

Hyperparameter is algorithm’s properties that set before running an ML algorithm.

**13. What do model-based learning algorithms search for? What is the most common strategy they use to succeed? How do they make predictions?**

Model-based learning search for groups and patterns and each group contains it’s particular properties.

**14. Can you name four of the main challenges in Machine Learning?**

Not enough data, Noisy data, wrong data and complex data.

**15. If your model performs great on the training data but generalizes poorly to new instances, what is happening? Can you name three possible solutions?**

Add more training data, simple the data and feature-engineering.

**16. What is a test set, and why would you want to use it?**

Test set used to check the performance of the model before deploying it to real life problem to check how well is generalize to actual problem.

**17. What is the purpose of a validation set?**

To check whether the model is working fine and how accurately it is predicting.

**18. What is the train-dev set, when do you need it, and how do you use it?**

Train-dev is validation set. Used to check how well model performed.

**19. What can go wrong if you tune hyperparameters using the test set?**

The model would be overfit.