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

Machine learning is a subset of Artificial intelligence. Machine learning is the ability to learn without explicitly being programmed.

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

Pattern recognizing, NLP, Synthetic environment, Anomaly detection

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

Data tagged with labels that characteristics or classifications or contained objects. It is used to set up supervised machine learning.

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

Regression and classification

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

Unsupervised learning uses clusters of the unlabelled dataset, and these data find patterns

or grouping without human help. Recommender systems, anomaly detection,

**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?**

K-means

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

Supervised Learning. Because the model should have some labeled data so that it can differentiate between spam and not spam.

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

Real-time data is used instead of normal data. Real-time processing is done.

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

Out-of-core learning is basically uses external memory. It is very usefull for large data set.

Core learning is based on RAM memory and data should be small compare to out-of-core learning.

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

Instance-based learning or similarity measures

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

Values that a model can learn about, like weight and coefficient of linear independence, weight and biases of a neural network, basically they are used for predictions.

Hyperparameters are those that are defined by the users, best values are extracted for the model (defined by the trail and error method). For example learning rate of a neural network, K in knn 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?**

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

Data quality and quantity, overfitting and underfitting, scalability, model selection, and evaluation.

**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?**

The model is overfitted to the training set, to overcome this issues we increase more training data, regularization,

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

After making the model on the training set we need to test the model's accuracy so the test set is used for checking the model's accuracy.

**17. What is a validation set's purpose?**

Sometimes only 80 20 is used, 80 for training and 20 for testing, to validate our model we use one more term called validation, and we divide it into 70(training) 20(testing) 10(validation)

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

Basically, it is training/development/test sets or development sets is also called validation sets.

During model development, tuning hyperparameters and before model evaluation on test sets validation sets is used. For smaller datasets, we considered it as 60 20 20, and for bigger datasets dividing datasets into 70 20 10 for training test and validation

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

The model will be overfitted to the test set. Performance matrices will give the wrong output. Data leakage.