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

Ans: - Machine learning is a type of AI which allows computers to predict outcomes without being programmed so. It uses historical data as input to predict outcomes.

This is done by feeding data to the models and allowing them to find patterns and make prediction about new data. Following are the application where ML is used:

1. Image Recognition
2. Recommendation Engine
3. Speech and text recognition
4. Fraud Detection

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

Ans: - Following are 4 distinct types of issues where ML shines:

1. Fraud Detection
2. Recommendation Engine
3. Speech Recognition
4. Object Detection

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

Ans: - Labeled training set a dataset with tagging to understand the data efficiently. Tags may be name, type or number. Machine learning algorithms learns by looking at the labels and understanding the patterns with the help of these labels.

Labeled data is very much essential for accuracy of the models, if the labeled dataset is not accurate then the ML algorithm will not be able to learn accurately. Following are some of the types of labeled training set:

1. Single label training set
2. Multi label training set
3. Semi-supervised training set
4. Active label training set

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

Ans: - Classification and Regression are the two important supervised task

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

Following are the four examples of Unsupervised tasks: -

1. Clustering
2. Dimensionality reduction
3. Anomaly detection
4. Association rule mining

**6.State the machine learning model that would be best to make a robot walk through various unfamiliar terrains?**

Ans: - Reinforcement Learning

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

Ans: - K-means clustering

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

Ans: -Spam detection is supervised learning problem, as without giving set of examples for spam and ham it won’t be able to find pattern in data.

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

Ans: - Online learning system refers to a digital platform that enables users to access content for study as well as resources too. So that users can acquire knowledge and skills without attending physical classes. Key features of online learning system are:

1. Course materials
2. Interactivity
3. Flexibility

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

Ans: - Out-of-core learning is a technique which is used to train machine learning algorithm on data that is too large to fit into the system. This is done by feeding the data into the system min small batches or chunks. This technique is more complex to implement but it is also very useful as it is scalable as we can add more processing power and servers.

On the other hand, core learning is a traditional approach of machine learning, in which the entire dataset is fed into the system at once. It is easy to implement in compare to out-of-core learning but also it is less efficient than the former. This technique is also easy to debug whereas former is more difficult to debug as data is not always available in memory

**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 is the value learned by the model on training dataset, it is used to make prediction on test dataset.

Whereas hyperparameter is a value set by the user before training the model, it controls the learning process and affects the performance of the model.

Differences:

1. Model parameter is internal to the model whereas hyperparameter is external to the model.
2. Model parameter learned from the model whereas hyperparameter is set by the users.
3. Model parameter affects the prediction pf the model whereas hyperparameter affects the learning process of the 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 look for following criteria to build predictive model:

1. Accuracy
2. Complexity
3. Generalization
4. Interpretability
5. Robustness

Most popular method they use to achieve success is minimizing loss function, for which they commonly use gradient descent and Bayesian optimization.

Following method, they use to make prediction:

1. Linear Regression
2. Logistic Regression
3. Decision Tree
4. Support Vector Machine
5. Neural Networks

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

Ans: - Following are the four most important Machine Learning challenges:

1. Data quality and availability
2. Model Interpretability
3. Model Overfitting
4. Model Deployment and Maintenance

**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: - When model performs well on training data but fails to generalize the result it is called overfitting. Overfitting can be happened for a number of reasons, some of them are as below:

1. Model has too many parameters or, model is too complex.
2. Training data is not sufficient or too small
3. Training data is too noisy

There is also different option can be chosen, some of them are as follows:

1. Choose different simple Model
2. Use regularization
3. Keep the same model and increase training data size

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

Ans: - A test set is a machine learning data subset which is used to evaluate the performance of the model, it is important to use a test set as it allows you to assess how well your model will generalize to new data.

If we only evaluate our model on training data it may end up over-fitting the model, and this can lead to poor performance on real world dataset.

Following are some of the benefits of using test set:

1. Prevents Overfitting
2. More realistic performance of model
3. Allows to compare more models with the developed one

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

Ans: - A validation set is a subset of machine learning dataset which is used to tune the hyperparameter of a trained model. Hyperparameters are the parameters that control the learning process of a machine learning model, such as the learning rate and number of epochs.

Purpose of fine tuning the model is to find the values that give the best performance on the validation set.

The validation set is important as it allows you to test the model and fine tuning the hyperparameters of your model without overfitting to the training dataset.

Using a validation set is important to assess the performance of the trained model and tune it to give the best values.

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

Ans: - A train-dev kit is a dataset that is split into two subsets: a training set and a development set. The training set is used to train a machine learning model, and the development set is used to evaluate the trained model and tune its hyperparameters.

The development set is also known as the validation set or the dev set. It is important to use a development set because it allows you to assess how well your model generalizes to new data, which is more realistic of the real-world conditions in which the model will be used.

If you only evaluate your model on the training data, you may end up overfitting the model. Overfitting occurs when the model learns the training data too well and is unable to generalize to new data. This can lead to poor performance on real-world tasks.

Here are some examples of situations where you would need to use a train-dev kit:

1. When you are developing a new machine learning model.
2. When you are improving an existing machine learning model.
3. When you are evaluating the performance of a machine learning model on new data.

Overall, using a train-dev kit is an important part of the machine learning process. It allows you to assess the performance of your model on unseen data and to prevent overfitting.

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

Ans: - If you use the test set to tune hyperparameters the model may overfit the test set. The test set is meant to be used to evaluate the performance of the trained model on new data.

Another risk of using the test set to tune hyperparameters is that you may run out of the data to evaluate the model. If you evaluate the final model on the test set, you will not be getting a true measure of how well the model will generalize to new data.