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

Ans - Machine learning refers to a field of artificial intelligence (AI) that focuses on the development of algorithms and models that enable computer systems to learn from and make predictions or decisions based on data, without being explicitly programmed for specific tasks. The core idea behind machine learning is to create systems that can automatically learn and improve from experience or data.

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

Ans – Credit card fault detection, Diamond price prediction, rain prediction, health care(if a patient is a cancer patient or not)

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

Ans - each data instance is associated with a corresponding label or target output. It serves as the input for training a machine learning model to learn the relationship between the input data and the corresponding labels.

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

Ans – Classification and Regression.

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

Ans - Clustering, Dimensionality Reduction, Anomaly Detection, Association Rule Mining

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

Ans -

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

Ans – k means clustering

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

Ans – supervised machine learning problem, we do not need any group or clustering over here.

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

Ans - The concept of an online learning system refers to a machine learning approach where the model learns and adapts continuously from incoming data in a dynamic and incremental manner. Unlike traditional batch learning, where the model is trained on a fixed dataset offline, online learning allows for real-time updates and immediate incorporation of new data for model training and prediction.

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

Ans - an approach in machine learning that deals with datasets that are too large to fit into the available memory (RAM) of a single machine. It allows for learning from data that is stored on disk or in an external storage system, enabling the processing of massive datasets that exceed the memory capacity.

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

Ans - The learning algorithm that makes predictions using a similarity measure is called Instance-based Learning or Lazy Learning.

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

Ans – Model parameter is the parameter which we used to give when we are training our model. But hyperparameter is the parameter which we will be needing in the time of hyper parameter tuning or in case of reducing the overfitting and to make our model’s accuracy much better.

1. 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 algorithms look for patterns and relationships in the training data to build a model or hypothesis that can make predictions or decisions.

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

Ans – Splitting the train and test data, fitting in transforming the data in such a way where data leakage will not take place, Overfitting and Underfitting, Feature selection and engineering.

1. 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 – It might be a situation of overfitting

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

Ans – Test set is the test data by which we will test how our model is predicting the output.

1. What is a validation set's purpose?

Ans – Validation set is nothing but the test set, which is require to test our model accuracy after training it by the training dataset. It is the unseen data for the model and by this we will see the performance of the model.

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

Ans - The train-dev set is a subset of the training data that is further split from the original training set. This split is created to help in model development, fine-tuning, and evaluation during the iterative process of model improvement.

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

Ans – It may cause overfitting to the test set, Biased performance estimation etc.