Viva Questions on ML

1. What do you understand by Machine learning?

Machine learning is the form of Artificial Intelligence that deals with system programming and automates data analysis to enable computers to learn and act through experiences without being explicitly programmed.For example, Robots are coded in such a way that they can perform the tasks based on data they collect from sensors. They automatically learn programs from data and improve with experiences.

2. Differentiate supervised and unsupervised machine learning.

In supervised machine learning, the machine is trained using labeled data. Then a new dataset is given into the learning model so that the algorithm provides a positive outcome by analyzing the labeled data. For example, we first require to label the data which is necessary to train the model while performing classification. In the unsupervised machine learning, the machine is not trained using labeled data and let the algorithms make the decisions without any corresponding output variables.

3. What is Supervised Learning?

Supervised learning is a machine learning algorithm of inferring a function from labeled training data. The training data consists of a set of training examples.

Example 1: Knowing the height and weight identifying the gender of the person.

Below are the popular supervised learning algorithms.

Support Vector Machines

Regression

Naive Bayes

Decision Trees

K-nearest Neighbour Algorithm and Neural Networks.

Example 2: If you build a T-shirt classifier, the labels will be "this is an S, this is an M and this is L", based on showing the classifier examples of S, M, and L.

4. Explain the terms Artificial Intelligence (AI), Machine Learning (ML) and Deep Learning?Artificial Intelligence (AI) is the domain of producing intelligent machines. ML refers to systems that can assimilate from experience (training data) and Deep Learning (DL) states to systems that learn from experience on large data sets. ML can be considered as a subset of AI. Deep
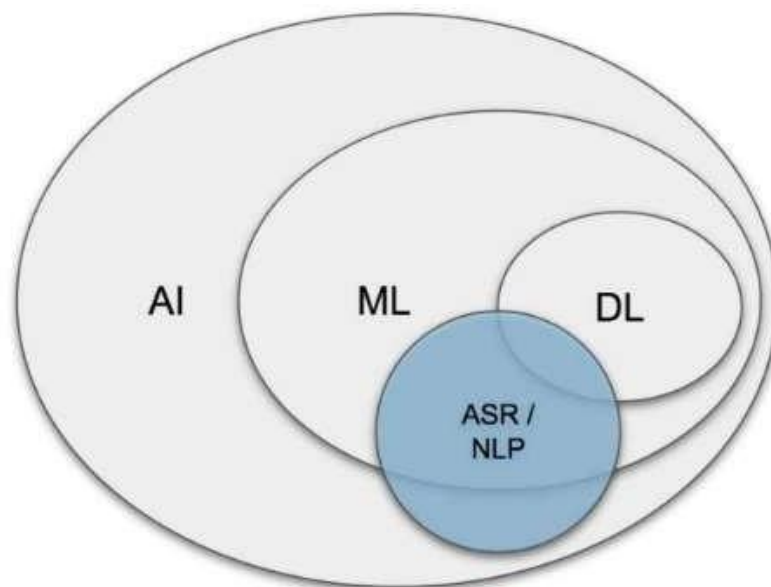
Learning (DL) is ML but useful to large data sets. The figure below roughly encapsulates the relation between AI, ML, and
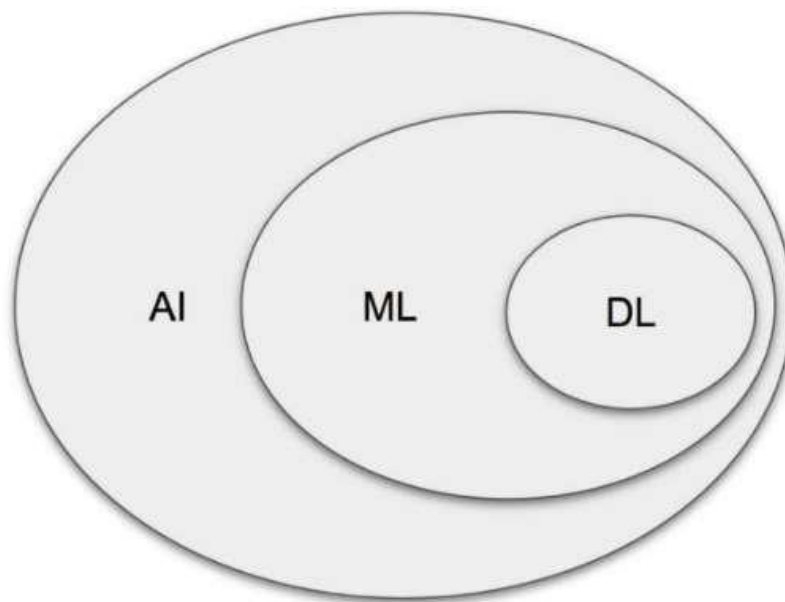
DL:

In summary, DL is a subset of ML & both were the subsets of AI.

Additional Information: ASR (Automatic Speech Recognition) & NLP (Natural Language Processing) fall under AI and overlay with ML & DL as ML is often utilized for NLP and ASR

tasks.

2. What are the different types of Learning/ Training models in ML?

ML algorithms can be primarily classified depending on the presence/absence of target variables.

A. *Supervised learning*: [Target is present]
The machine learns using labelled data. The model is trained on an existing data set before it starts making decisions with the new data.
The target variable is continuous: Linear Regression, polynomial Regression, and quadratic Regression.
The target variable is categorical: Logistic regression, Naive Bayes, KNN, SVM, Decision Tree, Gradient Boosting, ADA boosting, Bagging, Random forest etc.

B. *Unsupervised learning:* [Target is absent]
The machine is trained on unlabelled data and without any proper guidance. It automatically infers patterns and relationships in the data by creating clusters. The model learns through observations and deduced structures in the data.
Principal component Analysis, Factor analysis, Singular Value Decomposition etc.

C. *Reinforcement Learning:*
The model learns through a trial and error method. This kind of learning involves an agent that will interact with the environment to create actions and then discover errors or rewards of that action.

3. What is the difference between deep learning and machine learning?

Machine Learning involves algorithms that learn from patterns of data and then apply it to decision making. Deep Learning, on the other hand, is able to learn through processing data on its own and is quite similar to the human brain where it identifies something, analyse it, and makes a decision.
The key differences are as follows:

- The manner in which data is presented to the system.

- Machine learning algorithms always require structured data and deep learning networks rely on layers of artificial neural networks.
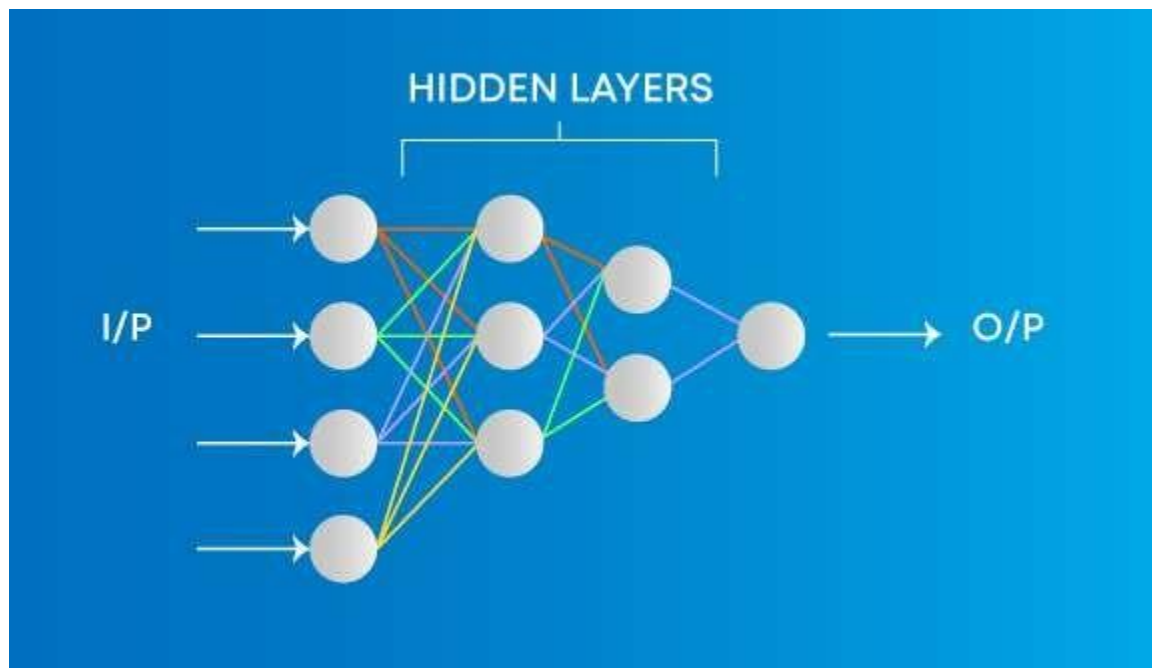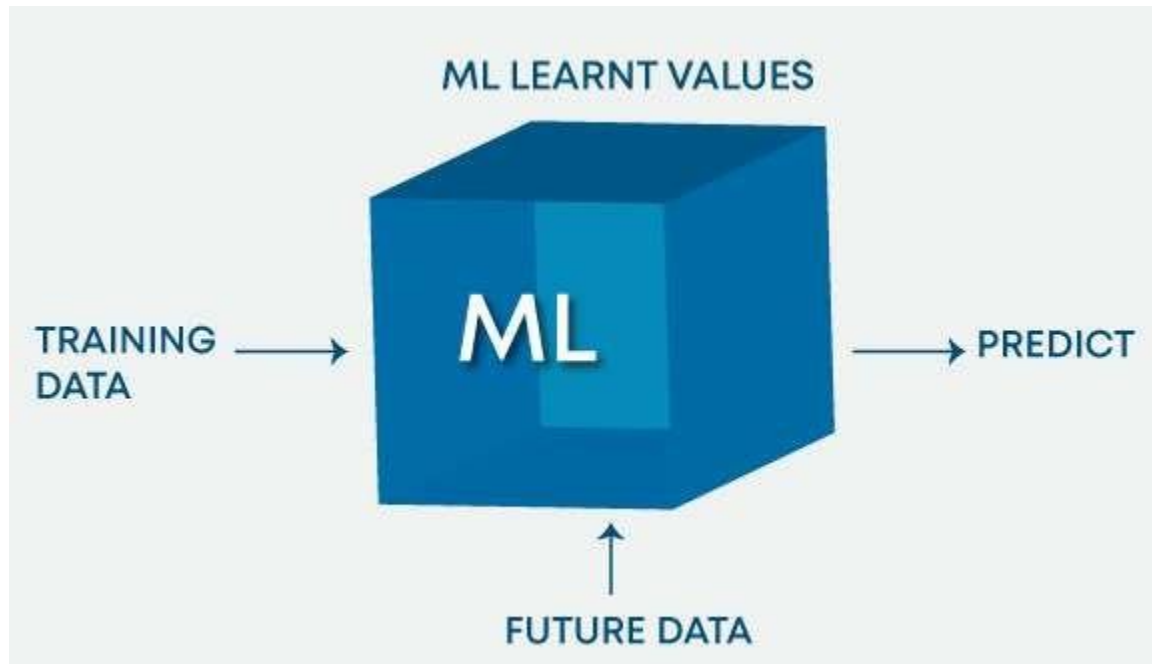
4. What is the main key difference between supervised and unsupervised machine learning?

Supervised learning	Unsupervised learning
The supervised learning technique needs labelled data to train the model. For example, to solve a classification problem (a supervised learning task), you need to have label data to train the model and to classify the data into your labelled groups.	Unsupervised learning does not need any labelled dataset. This is the main key difference between supervised learning and unsupervised learning.

5. How do you select important variables while working on a data set?

There are various means to select important variables from a data set that include the following:

- Identify and discard correlated variables before finalizing on important variables
- The variables could be selected based on 'p' values from Linear Regression
- Forward, Backward, and Stepwise selection
- Lasso Regression
- Random Forest and plot variable chart
- Top features can be selected based on information gain for the available set of features.

6. There are many machine learning algorithms till now. If given a data set, how can one determine which algorithm to be used for that?

Machine Learning algorithm to be used purely depends on the type of data in a given dataset. If data is linear then, we use linear regression. If data shows non-linearity then, the bagging algorithm would do better. If the data is to be analyzed/interpreted for some business purposes then we can use decision

trees or SVM. If the dataset consists of images, videos, audios then, neural networks would be helpful to get the solution accurately.

So, there is no certain metric to decide which algorithm to be used for a given situation or a data set. We need to explore the data using EDA (Exploratory Data Analysis) and understand the purpose of using the dataset to come up with the best fit algorithm. So, it is important to study all the algorithms in detail.

```
                          ┌──────────────────────┐
                          │  Machine Learning     │
                          └──────────────────────┘
```

Machine Learning

Supervised Learning

Unsupervised Learning

Reinforcement Learning

Classification/ Regression/ Estimation

Clustering/ Prediction

Decision-making

- Neural Networks.
- Support vector machine.
- Bayesian networks

- K means
- Gaussian mixture model
- Dirichlet process
- mixture model

- Q-learning
- R-learning
- TD learning