**1)Explain the following with an example:**

**1) Artificial Inteligence**

**2) Machine Learning**

**3) Deep Learning**

Ans- Artificial intelligence is a field of science concerned with building computers and machines that can reason, learn, and act in such a way that would normally require human intelligence or that involves data whose scale exceeds what humans can analyze.

Machine learning is a branch of [artificial intelligence (AI)](https://www.ibm.com/in-en/topics/artificial-intelligence) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.

Deep learning is a method in artificial intelligence (AI) that teaches computers to process data in a way that is inspired by the human brain. Deep learning models can recognize complex patterns in pictures, text, sounds, and other data to produce accurate insights and predictions.

**2) What is supervised learning? List some examples of supervised learning**.

Ans-  Supervised learning, also known as supervised machine learning, is a subcategory of machine learning and artificial intelligence. It is defined by its use of labeled datasets to train algorithms that to classify data or predict outcomes accurately.

The most commonly used Supervised Learning algorithms are decision tree, logistic regression, linear regression, support vector machine.

**3) What is unsupervised learning? List some examples of unsupervised learning**.

Ans- Unsupervised learning, also known as unsupervised machine learning, uses machine learning algorithms to analyze and cluster unlabeled datasets. These algorithms discover hidden patterns or data groupings without the need for human intervention.

Some examples of unsupervised learning algorithms include K-Means Clustering, Principal Component Analysis and Hierarchical Clustering.

**4) What is the difference between AI, ML, DL, and DS?**

Ans- **Artificial Intelligence is to make computers, to simulate the kind of things that humans can do (like playing chess or driving a car**and**many more), be able to solve such problems in ultimately better and faster way than people can do it.**

**Machine Learning is to make computers programmed at some given task, which is expected to learn from its environment and improve its performance over time.**

**There are a few complicated tasks, which are difficult to be solved using statistical models with machine learning. Deep Learning can be used to solve such challenging tasks in a much faster way with help of GPUs**.

**Data Science combines Machine Learning with Data Visualization, Data pre-processing, and also with other disciplines like Big Data Analytics and Cloud computing and Business knowledge.**

**5)** **What are the main differences between supervised, unsupervised, and semi-supervised learning?**

**Ans-** Supervised learning is the process where you have input variables (X) and an output variable (Y) and you use an algorithm to learn the mapping function from the input to the output.  
Y = f(X)

Supervised learning problems further grouped into regression and classification problems.  
**Classification**: A classification problem is when the output variable is a category, such as red or blue or disease and no disease.  
**Regression**: A regression problem is the output variable is a real value, such as dollars or weight.

Unsupervised learning is a machine learning technique, where you do not need to supervise the model. Instead, you need to allow the model to work on its own to discover information. It mainly deals with the unlabeled data. Algorithms are left to their own devises to discover and present the interesting structure in the data. Unsupervised learning is very useful in exploratory analysis because it can automatically identify structure in data.

Unsupervised learning problems can be further grouped into clustering and association problems.  
**Clustering**: Clustering is an important concept when it comes to unsupervised learning. It mainly deals with finding a structure or pattern in a collection of uncategorized data. Clustering algorithms will process your data and find natural clusters(groups) if they exist in the data.  
**Association**: An association rule learning problem is where you want to discover rules that describe large portions of your data. This unsupervised technique is about discovering exciting relationships between variables in large databases.

In Semi-supervised learning, the algorithm is trained upon a combination of labeled and unlabeled data. Typically, this combination will contain a very small amount of labeled data and a very large amount of unlabeled data.

**6) What is train, test and validation split? Explain the importance of each term.**

Ans- The difference between the training set and the validation set is the training set is the largest corpus of your dataset that you reserve for training your model. After training, inference on these images will be taken with a grain of salt, since the model has already had a chance to look at and memorize the correct output.

After all of the training experiments have concluded, you probably have gotten a sense on how your model might do on the validation set. But it is important to remember that the validation set metrics may have influenced you during the creation of the model, and in this sense you might, as a designer, overfit the new model to the validation set.

The validation split is a separate section of your dataset that you will use during training to get a sense of how well your model is doing on images that are not being used in training.

**7) How can unsupervised learning be used in anomaly detection?**

Ans- Unsupervised is useful when the data is unlabeled and thus requires less work to process the data set before- hand. Anomaly detection techniques of unsupervised nature assume that normal data points occur considerably more frequently than anomalous data points.

**8) List down some commonly used supervised learning algorithms and unsupervised learning algorithms.**

Ans- The most commonly used Supervised Learning algorithms are decision tree, logistic regression, linear regression, support vector machine. The most commonly used Unsupervised Learning algorithms are k-means clustering, hierarchical clustering, and apriori algorithm.