1. Explain the term machine learning, and how does it work? Explain two machine learning applications in the business world. What are some of the ethical concerns that machine learning applications could raise?

sol: So, machine learning is all about finding a pattern in the data, using it we can even predict the future if we have the required data specific to that domain.In the business we can predict the tax of the company, the required production so minimize the waste and many more.some of the ethical concerns are disappreance of jobs, lack of trust and the Cost to innovation.

2. Describe the process of human learning:

i. Under the supervision of experts

we can take the example of cycling, that first we see how people so cycling and then day by day we try under the supervision and finally crack the technique required for this.

ii. With the assistance of experts in an indirect manner

we can also learn in the assistance of experts like taking an online course and practicing what they are asking for us.

iii. Self-education

SElf\_education is learning from our own failures ,again and again trying to extract the desired results.

3. Provide a few examples of various types of machine learning.

sol: some of the types are:

1. Supervised learning
2. logistic regression
3. linear regression
4. decision tree
5. k-nearest neighbor
6. ensemble techniques
7. PCA
8. SVM
9. Unsupervised learning

a.clusterning(kMeans , DBSCAN)

1. Reinforcement learning

4. Examine the various forms of machine learning.

sol: In supervised learning, the output(label) is given, and using the required features we try to predict them whereas in unsupervised learning we give only the data and the machine tries to cluster them on the basis of their similarities. In reinforcement learning, we use the technique of reward and penalty.

5. Can you explain what a well-posed learning problem is? Explain the main characteristics that must be present to identify a learning problem properly.

sol: Well-posed learning problem means that a solution exists for that particular problem. the main characters are that all the required data is provided, performance measures must be known, and the learning task.

6. Is machine learning capable of solving all problems? Give a detailed explanation of your answer.

sol: “Most” of the problems can be solved by machine learning algos but you can’t say all of them, because automating each and every task is not possible. We can also think of it as like we cannot select the result of a children's music competition using it or we cannot predict the appraisal selecting any ml algo because we cannot feed the whole behavior of any person.

7. What are the various methods and technologies for solving machine learning problems? Any two of them should be defined in detail.

sol: logistic regression: This algo is generally used for the task of binary classification whose output is either 0 or 1, we can also use it for multi-classification and on the basis of features we can predict the output.

reinforcement learning: Reinforcement learning enables a machine to interact with an environment. A simple example is repeatedly playing a video game, and providing a reward when the algorithm takes the desired action. By repeating the process thousands or millions of times, the machine can eventually learn from its experience.

8. Can you explain the various forms of supervised learning? Explain each one with an example application.

### 1. Regression: In regression, a single output value is produced using training data. This value is a probabilistic interpretation, which is ascertained after considering the strength of correlation among the input variables. For example, regression can help predict the price of a house based on its locality, size, etc.

### 2. Classification:It involves grouping the data into classes. If you are thinking of extending credit to a person, you can use classification to determine whether or not a person would be a loan defaulter. When the supervised learning algorithm labels input data into two distinct classes, it is called binary classification. Multiple classifications means categorizing data into more than two classes.

### 3. Naive Bayesian Model:The Bayesian model of classification is used for large finite datasets. It is a method of assigning class labels using a direct acyclic graph. The graph comprises one parent node and multiple children nodes. And each child node is assumed to be independent and separate from the parent.

Decision Trees: A decision tree is a flowchart-like model that contains conditional control statements, comprising decisions and their probable consequences. The output relates to the labelling of unforeseen data.

### 4. Random Forest Model: The random forest model is an ensemble method. It operates by constructing a multitude of decision trees and outputs a classification of the individual trees. Suppose you want to predict which undergraduate students will perform well in GMAT – a test taken for admission into graduate management programs. A random forest model would accomplish the task, given the demographic and educational factors of a set of students who have previously taken the test.

9. What is the difference between supervised and unsupervised learning? With a sample application in each region, explain the differences.

sol: The chief and the only reason between these two are that the label is provided in the supervised learning that is the output is already given and from the attributes and the label we try to make machine learn the pattern for example if we want to predict the annual income, we have to feed the machine some data with the label as the annual income. Whereas in unsupervised learning the machine itself try to label each cluster it has made on the basis of similarities they are having, on this basis customer segmentation is made.

10. Describe the machine learning process in depth.

sol: So in every machine learning process first we try to get the depth knowledge of the problem, then we try to gather the data required.After performing some of the data cleaning techniques we try to make a structured data. Then we perform some of EDA tasks so that we can go for feature selection and feature extraction part. And after that we go for seelcting the best algos for that particular problem.Then comes the validation part which gives the performance of the model.

a. Make brief notes on any two of the following:

MATLAB is one of the most widely used programming languages.

ii. Deep learning applications in healthcare

sol: So in healthcare, DL can maximize the outcome by giving the X-rays we predict if cancel is detected or not. first, we will gather some previous data, then after feeding all that data to the machine(algo) we will make a prediction model which by just seeing the X-ray report can predict that if the cancer is there or not.

iii. Study of the market basket

iv. Linear regression (simple)

sol: Linear regression is a supervised learning algorithm in which we try to find the best fitted line so that we can predict the outcomes in future. we just have to decrease the cost in order to find the slope and y intercept which is required to get a line.

11. Make a comparison between:-

1. Generalization and abstraction

sol: Generalization is like that machine is generalizing or categorizing what is what whereas in case of abstraction machine is giving the output like the predicted price of the house.

2. Learning that is guided and unsupervised

sol: The chief and the only reason between these two are that the label is provided in the supervised learning that is the output is already given and from the attributes and the label we try to make a machine learn the pattern for example if we want to predict the annual income, we have to feed the machine some data with the label as the annual income. Whereas in unsupervised learning the machine itself try to label each cluster it has made on the basis of similarities they are having, on this basis customer segmentation is made.

3. Regression and classification

sol: The main difference between Regression and Classification algorithms that Regression algorithms are used to predict the continuous values such as price, salary, age, etc., and Classification algorithms are used to predict/Classify the discrete values such as Male or Female, True or False, Spam or Not Spam, etc.