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?

**Ans**: machine learning is a core sub-area of Artificial Intelligence (AI). ML applications learn from experience (or to be accurate, data) like humans do without direct programming. When exposed to new data, these applications learn, grow, change, and develop by themselves.

The Machine Learning process starts with inputting training data into the selected algorithm. Training data being known or unknown data to develop the final Machine Learning algorithm. The type of training data input does impact the algorithm, and that concept will be covered further momentarily. New input data is fed into the machine learning algorithm to test whether the algorithm works correctly. The prediction and results are then checked against each other.If the prediction and results don’t match, the algorithm is re-trained multiple times until the data scientist gets the desired outcome. This enables the machine learning algorithm to continually learn on its own and produce the optimal answer, gradually increasing in accuracy over time. Application:

### 1.Customer Lifetime Value Modeling are imp for eCommerce business to employ:

That’s because they can be used to identify, understand, and retain your company's most valuable customers, whether that means the biggest spenders, the most loyal advocates of your brand, or both. These models predict the future revenue that an individual customer will bring to your business in a given period. With this information, you can focus your marketing efforts to encourage these customers to interact with your brand more often and even target your acquisition spend to attract new customers that are similar to your existing MVPs

### 2.Customer Segmentation: Rather than relying on a marketer’s intuition to separate customers into groups for campaigns, data scientists can use clustering and classification algorithms to group customers into personas based on specific variations among them. These personas account for customer differences across multiple dimensions such as demographics, browsing behavior, and affinity. Connecting these traits to patterns of purchasing behavior allows data-savvy companies to roll out highly personalized marketing campaigns that are more effective at boosting sales than generalized campaigns.

1. Describe the process of human learning:

Ans:

i. Under the supervision of experts:

Under supervision of experts ,experts are providing some sample may be toy of different shape ,color and sizes and human learns it and experts if ask for similar but other variety.then human needs to recognise it which is that toy/thing.

ii. With the assistance of experts in an indirect manner:experts are supporting the for identification but not on previous seen toys or things.

iii. Self-education:in self education human learns by itself either by geeting some revards/penalty. for example

1. Provide a few examples of various types of machine learning:

Ans:

* **Supervised learning** algorithms take labelled inputs and map them to the known outputs, which means you already know the target variable. e.g-weather prediction, sales forecasting, stock price analysis.
* **Unsupervised learning** is a type of machine learning that uses unlabelled data to train machines. Unlabeled data doesn’t have a fixed output variable. The model learns from the data, discovers the patterns and features in the data, and returns the output.e.g.= uses the images of vehicles to classify if it’s a bus or a truck., customer segmentation.
* **Reinforcement Learning**: trains a machine to take suitable actions and maximize its rewards in a particular situation. It uses an agent and an environment to produce actions and rewards.e.g gaming industries to build games

1. Examine the various forms of machine learning.

**Supervised learning** algorithms take labelled inputs and map them to the known outputs, which means you already know the target variable. e.g-weather prediction, sales forecasting, stock price analysis.

**Unsupervised learning** is a type of machine learning that uses unlabelled data to train machines. Unlabeled data doesn’t have a fixed output variable. The model learns from the data, discovers the patterns and features in the data, and returns the output.e.g.= uses the images of vehicles to classify if it’s a bus or a truck., customer segmentation.

**Reinforcement Learning**: trains a machine to take suitable actions and maximize its rewards in a particular situation. It uses an agent and an environment to produce actions and rewards.e.g gaming industries to build games

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

**Ans: Well Posed Learning Problem:**A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E. To have a well-defined learning problem, three features needs to be identified:

* + The class of tasks(T)
  + The measure of performance to be improved(P)
  + The source of experience(E)
* **Examples:**A checkers learning problem:
  + Task T: playing checkers.
  + Performance measure P: percent of games won against opponents.
  + Training experience E: playing practice games against itself.
  + A Self driving learning problem:
  + Task T: driving on public routes using vision sensors
  + Performance measure P: average distance travelled before an error (as judged by human overseer)
  + Training experience E: a sequence of images and steering commands recorded while observing a human driver.

Q6. Is machine learning capable of solving all problems? Give a detailed explanation of your answer.

Ans: The short answer is no. But, like any algorithm, Machine Learning is a tool.  It is a tool that can be used to help make decisions, classifications, and predictions in various domains, but obviously it is not omnipotent to all problems. Sometimes standard statistical methods are all that is required. Furthermore, how effectively AI/ML can solve problems also depends on the quality of data and data scientists.

Q7. What are the various methods and technologies for solving machine learning problems? Any two of them should be defined in detail.

Ans: Machine learning techniques can be divided into three broad categories: supervised, unsupervised, and reinforcement learning. Supervised learning algorithms are used to create a model based on known input and output data to make future predictions. Unsupervised learning detects hidden patterns in input data to make its predictions.Machine learning algorithms generally depend on [frameworks](https://www.upwork.com/resources/top-machine-learning-frameworks/), such as [**TensorFlow**](https://www.tensorflow.org/)and **[PyTorch](https://pytorch.org/" \t "_blank)**for Python, to speed up solutions development.

## **Linear Regression**

Regression algorithms fall under the umbrella of supervised learning–they predict solutions according to prior data. For example, a regression model could predict how a group of stocks might do based on historical data.There are many different kinds of regression techniques or methods, including polynomial regression, decision trees, neural nets, random forest regression, logistic regression, regularization, and the Gaussian process. But the simplest one, and the best place to start, is linear regression. A simple linear regression works out the relationship between at least two variables, which can be dependent (target variables) and independent (predictor) variables. A model set is created using the mathematical equation y = m \* x + b. Data pairs (x,y) are used for training data in a linear model and calculating the position and slope of a line, in which the total distance among all the data points is minimized.

## **Classification**

Classification is a supervised machine learning method that involves predicting and sometimes explaining a class value. With the right input data, for example, it might be able to predict whether a customer will purchase a certain product, with the possible outcomes being “yes buyer” or “no buyer.” Of course, this method can cover situations with more than two possible outcomes, such as predicting different credit scores.Classification works best when you can easily tag, categorize, or separate the data into specific groups or classes. A classified model analyzes input to learn how to classify new data. Different types of algorithms for classifiers include binary, multiclass, and multilabel.

Q8. Can you explain the various forms of supervised learning? Explain each one with an example application.

Ans: Various forms of ML: Supervised, Unsupervised and Reinforcement learning.Supervised learning algorithms take labelled inputs and map them to the known outputs, which means you already know the target variable. Supervised Learning methods need external supervision to train machine learning models. Hence, the name supervised. They need guidance and additional information to return the desired result.

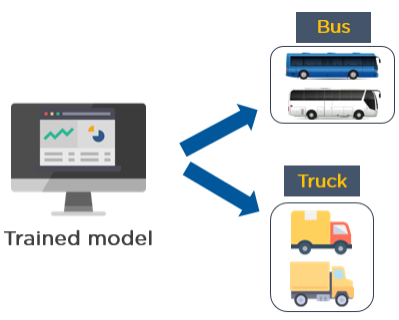
**Applications:**Supervised learning algorithms are generally used for solving classification and regression problems.



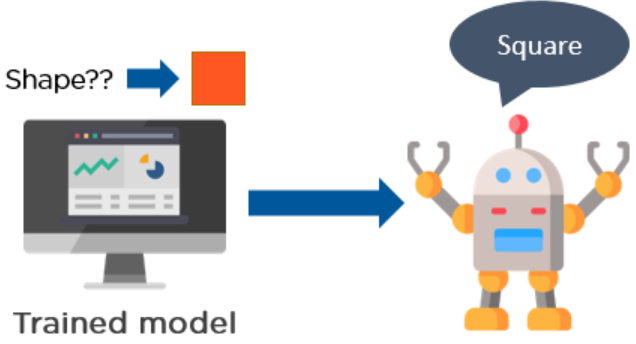
Few of the top supervised learning applications are weather prediction, sales forecasting, stock price analysis.

Q9. What is the difference between supervised and unsupervised learning? With a sample application in each region, explain the differences.

Ans: Unsupervised learning is a type of machine learning that uses unlabeled data to train machines. Unlabeled data doesn’t have a fixed output variable. The model learns from the data, discovers the patterns and features in the data, and returns the output.Depicted below is an example of an unsupervised learning technique that uses the images of vehicles to classify if it’s a bus or a truck. The model learns by identifying the parts of a vehicle, such as a length and width of the vehicle, the front, and rear end covers, roof hoods, the types of wheels used etc. Based on these features, the model classifies if the vehicle is a bus or a truck.



Reinforcement Learning: Reinforcement Learning trains a machine to take suitable actions and maximize its rewards in a particular situation. It uses an agent and an environment to produce actions and rewards. The agent has a start and an end state. But, there might be different paths for reaching the end state, like a maze. In this learning technique, there is no predefined target variable. An example of reinforcement learning is to train a machine that can identify the shape of an object, given a list of different objects. In the example shown, the model tries to predict the shape of the object, which is a square in this case.



### Algorithms: Q-learning,Sarsa,Monte Carlo ,Deep Q network

Q10. Describe the machine learning process in depth.

Ans: Machine learning starts with data — numbers, photos, or text, like bank transactions, pictures of people or even [bakery items](https://www.newyorker.com/tech/annals-of-technology/the-pastry-ai-that-learned-to-fight-cancer), repair records, time series data from sensors, or sales reports. The data is gathered and prepared to be used as training data, or the information the machine learning model will be trained on. The more data, the better the program. From there, programmers choose a machine learning model to use, supply the data, and let the computer model train itself to find patterns or make predictions. Over time the human programmer can also tweak the model, including changing its parameters, to help push it toward more accurate results. Some data is held out from the training data to be used as evaluation data, which tests how accurate the machine learning model is when it is shown new data. The result is a model that can be used in the future with different sets of data.

Q.11 Make brief notes on any two of the following:

Deep learning applications in healthcare

### Ans: Medical imaging , Healthcare data analytics .Mental health chatbots, Personalized medical treatments.Meidcal Insurance,Prescription audit, Responding to patient queries

1. Medical imaging: Image recognition and object detection are used in Magnetic Resonance (MR) and Computed tomography (CT) processes for image segmentation, disease detection & prediction. Deep learning models can make effective interpretations by a combination of aspects of imaging data, for example, tissue size, volume, and shape. These models can flag important areas in images. For example,deep learning algorithms are used for diabetic retinopathy detection, early detection of Alzheimer and ultrasound detection of breast nodules. Thanks to new advances in deep learning, most pathology and radiology images can be investigated in the future. Deep learning algorithms simplify complex data analysis, so abnormalities are determined and prioritized more precisely. The insights that convolutional neural networks (CNNs) provide, help medical professionals to notice the health issues of their patients on time and more accurately.

### 2. Healthcare data analytics: Deep learning models can analyze electronic health records (EHR) that contain structured and unstructured data, including clinical notes, laboratory test results, diagnosis, and medications at exceptional speeds with the most possible accuracy. Also, smartphones and[wearable devices provide useful information about lifestyle. They have the](https://www.businessinsider.com/wearable-technology-healthcare-medical-devices?r=US&IR=T) potential to transform data by using mobile apps to monitor medical risk factors for deep learning models. In 2019, Current Health’s AI wearable device [became](https://www.docwirenews.com/docwire-pick/future-of-medicine-picks/first-ai-wearable-approved-by-fda-for-home-use-monitoring-vitals/) one of the first AI medical monitoring wearables approved by Food and Drug Administration (FDA) for use at home. This device can measure the pulse, respiration, oxygen saturation, temperature, and mobility of patients.

### iv. Linear regression (simple):

### Ans: Simple linear is an approach for predicting the quantitative response Y based on single predictor variable X.

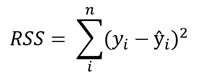


This is the equation of straight-line having slope β1 and intercept β0. Initially, we will consider the simple linear regression model for the sales and money spent on TV advertising media.Then the mathematical equation becomes **𝑆𝑎𝑙𝑒𝑠 = 𝛽0 + 𝛽1 \* 𝑇𝑉.**

Step 1: Estimating the coefficients: (Let’s find the coefficients)

Now to find the estimate of the sales for the advertising budget, we have to know the values of the β1 and β0. For the best estimate, the difference between predicted sales and the actual sales (called as **residual**) should be minimum.

As the residual may be negative or positive, so while calculating the net residual it can be lead to cancellation terms and reduction of net effect which leads to a non-optimal estimate of coefficients. To overcome this, we use a Residual sum of squares (RSS).



With a simple calculation, we can find the value of β0 and β1 for minimum RSS value.

Text

Description automatically generated

With the stats model library in python, we can find out the coefficients,

Shape

Description automatically generated

Table 1: Simple regression of sales on TVValues for β0 and β1 are 7.03 and 0.047 respectively. Then the relation becomes, **Sales = 7.03 + 0.047 \* TV.**T his means if we spend an additional 1000 dollars on TV advertising media it increases the sales of products by 47 units. *This gives us how strongly the TV advertising media associated with the sales.*Step 2: Assessing the Accuracy of the Coefficient Estimates ( How accurate these coefficients are? ) Standard Error (SE). SE of a coefficient represents the average distance that observed values deviate from the regression line. If the standard errors of the coefficient estimate of a variable are smaller then the model can estimate the coefficient for that variable with greater precision.SE of the coefficient of TV and intercept can be given by,



## The standard error is used to compute, Confidence Interval, To Perform Hypothesis Testing and Accuracy of Model

Q12. Make a comparison between: -

Ans: 1. Generalization and abstraction:

**abstraction operation**

* changes the representation of an entity by hiding/reducing its properties that are not necessary for the desired conceptualization
* has an inherent information loss, which makes it less flexible but more conclusive

**generalization operation**

* doesn't change the representation of an entity, but defines similarities between entities of different kind
* applies knowledge previously acquired to unseen circumstances or extends that knowledge beyond the scope of the original problem (knowledge transfer)
* can be seen as a hypothesis that a set of entities of different kind have similar properties and will behave consistently when applied in a certain way
* has no inherent information loss, which makes it more flexible but less conclusive (more error-prone)

2. Learning that is guided and unsupervised

Ans: [Supervised learning](https://www.ibm.com/cloud/learn/supervised-learning) is a machine learning approach that’s defined by its use of labeled datasets. These datasets are designed to train or “supervise” algorithms into classifying data or predicting outcomes accurately. Using labeled inputs and outputs, the model can measure its accuracy and learn over time.

[Unsupervised learning](https://www.ibm.com/cloud/learn/unsupervised-learning) uses machine learning algorithms to analyze and cluster unlabeled data sets. These algorithms discover hidden patterns in data without the need for human intervention (hence, they are “unsupervised”).

3. Regression and classification

Ans: **Classification** problems use an algorithm to accurately assign test data into specific categories, such as separating apples from oranges. Or, in the real world, supervised learning algorithms can be used to classify spam in a separate folder from your inbox. Linear classifiers, support vector machines, decision trees and [random forest](https://www.ibm.com/cloud/learn/random-forest) are all common types of classification algorithms.

**Regression** is another type of supervised learning method that uses an algorithm to understand the relationship between dependent and independent variables. Regression models are helpful for predicting numerical values based on different data points, such as sales revenue projections for a given business. Some popular regression algorithms are linear regression, logistic regression and polynomial regression.