**ASSIGNMENT NO. 01**

**Q.1] Define machine learning and explain difference between Traditional approach and Machine Learning approach.**

* **Machine learning :-** “Machine learning (ML) is defined as a discipline of artificial intelligence (AI) that provides machines the ability to automatically learn from data and past experiences to identify patterns and make predictions with minimal human intervention. ” The Capability of Artificial Intelligence Systems To Learn By Extracting Patterns From Data Known as Machine Learning.
* **Traditional Approach :** Traditional Programming relies on hard-coded rules.

|  |
| --- |
| https://techiethoughtss.files.wordpress.com/2020/01/traditional-app.png?w=1024 |

* **Machine learning Approach :** Machine learning relies on learning patterns based on sample data

|  |
| --- |
| mls3 0102 |

**Q.2] Describe Supervised machine learning technique with suitable example.**

* Supervised learning uses labeled data (data with known answers) to train algorithms to: **1] Classify Data. 2] Predict Outcomes.**
* Supervised learning can classify data like "What is spam in an e-mail", based on known spam examples.
* Supervised learning can predict outcomes like predicting what kind of video you like, based on the videos you have played.
* Supervised machine learning is based on supervision. It means in the supervised learning technique, we train the machines using the "labelled" dataset, and based on the training, the machine predicts the output.
* Here, the labelled data specifies that some of the inputs are already mapped to the output. More preciously, we can say; first, we train the machine with the input and corresponding output, and then we ask the machine to predict the output using the test dataset.
* **Example:** **Suppose we have an input dataset of cats and dog images. :-**
* So, first, we will provide the training to the machine to understand the images, such as the shape & size of the tail of cat and dog, Shape of eyes, colour, height (dogs are taller, cats are smaller), etc
* After completion of training, we input the picture of a cat and ask the machine to identify the object and predict the output. Now, the machine is well trained, so it will check all the features of the object, such as height, shape, colour, eyes, ears, tail, etc., and find that it's a cat. So, it will put it in the Cat category. This is the process of how the machine identifies the objects in Supervised Learning.
* The main goal of the supervised learning technique is to map the input variable(x) with the output variable(y).

**Q.3 Explain Unsupervised machine learning technique with suitable example.**

* Unsupervised learning is used to predict undefined relationships like meaningful patterns in data.
* It is about creating computer algorithms than can improve themselves.
* It is expected that machine learning will shift to unsupervised learning to allow programmers to solve problems without creating models.
* Unsupervised learning is different from the Supervised learning technique; as its name suggests, there is no need for supervision. It means, in unsupervised machine learning, the machine is trained using the unlabeled dataset, and the machine predicts the output without any supervision.
* In unsupervised learning, the models are trained with the data that is neither classified nor labelled, and the model acts on that data without any supervision.
* The main aim of the unsupervised learning algorithm is to group or categories the unsorted dataset according to the similarities, patterns, and differences. Machines are instructed to find the hidden patterns from the input dataset.
* **Example:** Suppose there is a basket of fruit images, and we input it into the machine learning model. The images are totally unknown to the model, and the task of the machine is to find the patterns and categories of the objects.
* So, now the machine will discover its patterns and differences, such as colour difference, shape difference, and predict the output when it is tested with the test dataset.

**Q.4] What are the advantages and disadvantages of Supervised machine learning?**

* **Advantages:**
* Since supervised learning work with the labelled dataset so we can have an exact idea about the classes of objects.
* These algorithms are helpful in predicting the output on the basis of prior experience.
* **Disadvantages:**
* These algorithms are not able to solve complex tasks.
* It may predict the wrong output if the test data is different from the training data.
* It requires lots of computational time to train the algorithm

**Q.5] What are the advantages and disadvantages of Unsupervised machine learning?**

* **Advantages:**
* These algorithms can be used for complicated tasks compared to the supervised ones because these algorithms work on the unlabeled dataset.
* Unsupervised algorithms are preferable for various tasks as getting the unlabeled dataset is easier as compared to the labelled dataset.
* **Disadvantages:**
* The output of an unsupervised algorithm can be less accurate as the dataset is not labelled, and algorithms are not trained with the exact output in prior.
* Working with Unsupervised learning is more difficult as it works with the unlabelled dataset that does not map with the output.

**Q.6] Define reinforcement learning and explain how Reinforcement learning works?**

* Reinforcement learning works on a feedback-based process, in which an AI agent (A software component) automatically explore its surrounding by hitting & trail, taking action, learning from experiences, and improving its performance. Agent Gets rewarded for each good action and get punished for each bad action; hence the goal of reinforcement learning agent is to maximize the rewards.
* In reinforcement learning, there is no labelled data like supervised learning, and agents learn from their experiences only.
* The Reinforcement learning process is similar to a human being;
* **Example:** A child learns various things by experiences in his day-to-day life. An example of reinforcement learning is to play a game, where the Game is the environment, moves of an agent at each step define states, and the goal of the agent is to get a high score. Agent receives feedback in terms of punishment and rewards.
* Due to its way of working, reinforcement learning is employed in different fields such as Game theory, Operation Research, Information theory, multi-agent systems.
* **Advantages:**
* It helps in solving complex real-world problems which are difficult to be solved by general techniques.
* The learning model of RL is similar to the learning of human beings; hence most accurate results can be found.
* Helps in achieving long term results.
* **Disadvantage:**
* RL algorithms are not preferred for simple problems.
* RL algorithms require huge data and computations.
* Too much reinforcement learning can lead to an overload of states which can weaken the results

**Q.7] Explain Classification problem with suitable example.**

* As we know, the Supervised Machine Learning algorithm can be broadly classified into Regression and Classification Algorithms. In Regression algorithms, we have predicted the output for continuous values, but to predict the categorical values, we need Classification algorithms.
* The Classification algorithm is a Supervised Learning technique that is used to identify the category of new observations on the basis of training data.
* In Classification, a program learns from the given dataset or observations and then classifies new observation into a number of classes or groups. Such as, Yes or No, 0 or 1, Spam or Not Spam, cat or dog, etc. Classes can be called as targets/labels or categories.
* Unlike regression, the output variable of Classification is a category, not a value, such as "Green or Blue", "fruit or animal", etc. Since the Classification algorithm is a Supervised learning technique, hence it takes labeled input data, which means it contains input with the corresponding output.
* In classification algorithm, a discrete output function(y) is mapped to input variable(x),

y=f(x), where y = categorical output

* The best example of an ML classification algorithm is Email Spam Detector.
* The main goal of the Classification algorithm is to identify the category of a given dataset, and these algorithms are mainly used to predict the output for the categorical data.
* Classification algorithms can be better understood using the below diagram. In the below diagram, there are two classes, class A and Class B. These classes have features that are similar to each other and dissimilar to other classes.
* The algorithm which implements the classification on a dataset is known as a classifier. There are two types of Classifications:
* **Binary Classifier:** If the classification problem has only two possible outcomes, then it is called as Binary Classifier.

**Examples:** YES or NO, MALE or FEMALE, SPAM or NOT SPAM, CAT or DOG, etc.

* **Multi-class Classifier:** If a classification problem has more than two outcomes, then it is called as Multi-class Classifier.

**Example:** Classifications of types of crops, Classification of types of music.

**Q.8] Explain Regression problem with suitable example.**

* As we know, the Supervised Machine Learning algorithm can be broadly classified into
* Regression and
* Classification Algorithms.
* Regression analysis is a statistical method to model the relationship between a dependent (target) and independent (predictor) variables with one or more independent variables.
* More specifically, Regression analysis helps us to understand how the value of the dependent variable is changing corresponding to an independent variable when other independent variables are held fixed. It predicts continuous/real values such as temperature, age, salary, price, etc.
* **Example:** Suppose there is a marketing company A, who does various advertisement every year and get sales on that. The below list shows the advertisement made by the company in the last 5 years and the corresponding sales:
* Now, the company wants to do the advertisement of $200 in the year 2019 and wants to know the prediction about the sales for this year. So to solve such type of prediction problems in machine learning, we need regression analysis.
* Regression is a supervised learning technique which helps in finding the correlation between variables and enables us to predict the continuous output variable based on the one or more predictor variables.
* It is mainly used for prediction, forecasting, time series modeling, and determining the causal-effect relationship between variables.
* Some examples of regression can be as:
  + Prediction of rain using temperature and other factors
  + Determining Market trends
  + Prediction of road accidents due to rash driving.

**Q.9] Explain Clustering problem with suitable example.**

* Clustering or cluster analysis is a machine learning technique, which groups the unlabelled dataset. It can be defined as "A way of grouping the data points into different clusters, consisting of similar data points. The objects with the possible similarities remain in a group that has less or no similarities with another group."
* It does it by finding some similar patterns in the unlabelled dataset such as shape, size, color, behavior, etc., and divides them as per the presence and absence of those similar patterns.
* It is an unsupervised learning method, hence no supervision is provided to the algorithm, and it deals with the unlabeled dataset.
* After applying this clustering technique, each cluster or group is provided with a cluster-ID. ML system can use this id to simplify the processing of large and complex datasets.
* **Example:** When we visit any shopping mall, we can observe that the things with similar usage are grouped together. Such as the t-shirts are grouped in one section, and trousers are at other sections, similarly, at vegetable sections, apples, bananas, Mangoes, etc., are grouped in separate sections, so that we can easily find out the things. • The clustering technique also works in the same way.
* **Example:** Apart from these general usages, it is used by the Amazon in its recommendation system to provide the recommendations as per the past search of products. • Netflix also uses this technique to recommend the movies and web-series to its users as per the watch history. • The below diagram explains the working of the clustering algorithm. We can see the different fruits are divided into several groups with similar properties.
* The clustering technique can be widely used in various tasks. Some most common uses of this technique are:
* Market Segmentation
* Statistical data analysis
* Social network analysis
* Image segmentation
* Anomaly detection, etc.

**Q.10] Diffentiate between data, information and knowledge.**

* **Data :** Data, information, and knowledge are often used interchangeably. However, these terms represent different stages of value creation from data to decision-making. Data are the raw alphanumeric values obtained through different acquisition methods. Data in their simplest form consist of raw alphanumeric values.
* **Information :** Information is created when data are processed, organized, or structured to provide context and meaning. Information is essentially processed data. Knowledge is what we know.
* **Knowledge :** Knowledge is unique to each individual and is the accumulation of past experience and insight that shapes the lens by which we interpret, and assign meaning to, information. For knowledge to result in action, an individual must have the authority and capacity to make and implement a decision. Knowledge (and authority) are needed to produce actionable information that can lead to impact.

**Q.11] Describe Train Test Split in machine learning.**

* In Machine Learning we create models to predict the outcome of certain events, like in the previous chapter where we predicted the CO2 emission of a car when we knew the weight and engine size.
* To measure if the model is good enough, we can use a method called Train/Test.
* Train/Test is a method to measure the accuracy of your model. It is called Train/Test because you split the the data set into two sets: a training set and a testing set. 80% for training, and 20% for testing. You train the model using the training set. You test the model using the testing set. Train the model means create the model. Test the model means test the accuracy of the model.

**Q.12] Short short note on “Applications of machine learning”.**

* Machine learning is a buzzword for today's technology, and it is growing very rapidly day by day. We are using machine learning in our daily life even without knowing it such as Google Maps, Google assistant, Alexa, etc. Below are some most trending real-world applications of Machine Learning:
* **A] Image Recognition:** • Image recognition is one of the most common applications of machine learning. It is used to identify objects, persons, places, digital images, etc. The popular use case of image recognition and face detection is, Automatic friend tagging suggestion: • Facebook provides us a feature of auto friend tagging suggestion. Whenever we upload a photo with our Facebook friends, then we automatically get a tagging suggestion with name, and the technology behind this is machine learning's face detection and recognition algorithm. • It is based on the Facebook project named "Deep Face," which is responsible for face recognition and person identification in the picture.
* **B] Speech Recognition** • While using Google, we get an option of "Search by voice," it comes under speech recognition, and it's a popular application of machine learning. • Speech recognition is a process of converting voice instructions into text, and it is also known as "Speech to text", or "Computer speech recognition." • At present, machine learning algorithms are widely used by various applications of speech recognition. Google assistant, Siri, Cortana, and Alexa are using speech recognition technology to follow the voice instructions.
* **C] Traffic prediction:** • If we want to visit a new place, we take help of Google Maps, which shows us the correct path with the shortest route and predicts the traffic conditions. • It predicts the traffic conditions such as whether traffic is cleared, slow-moving, or heavily congested with the help of two ways: • Real Time location of the vehicle form Google Map app and sensors • Average time has taken on past days at the same time. • Everyone who is using Google Map is helping this app to make it better. It takes information from the user and sends back to its database to improve the performance.
* **D] Product Recommendations:** • Machine learning is widely used by various e-commerce and entertainment companies such as Amazon, Netflix, etc., for product recommendation to the user. Whenever we search for some product on Amazon, then we started getting an advertisement for the same product while internet surfing on the same browser and this is because of machine learning. • Google understands the user interest using various machine learning algorithms and suggests the product as per customer interest. • As similar, when we use Netflix, we find some recommendations for entertainment series, movies, etc., and this is also done with the help of machine learning.
* **E] Self-driving cars:** • One of the most exciting applications of machine learning is self-driving cars. Machine learning plays a significant role in self-driving cars. Tesla, the most popular car manufacturing company is working on self-driving car. It is using unsupervised learning method to train the car models to detect people and objects while driving.
* **F] Email Spam and Malware Filtering:** • Whenever we receive a new email, it is filtered automatically as important, normal, and spam. We always receive an important mail in our inbox with the important symbol and spam emails in our spam box, and the technology behind this is Machine learning. Below are some spam filters used by Gmail: • Content Filter • Header filter • General blacklists filter • Rules-based filters • Permission filters.
* **G] Virtual Personal Assistant:** • We have various virtual personal assistants such as Google assistant, Alexa, Cortana, Siri. • As the name suggests, they help us in finding the information using our voice instruction. • These assistants can help us in various ways just by our voice instructions such as Play music, call someone, Open an email, Scheduling an appointment, etc. • These virtual assistants use machine learning algorithms as an important part. • These assistant record our voice instructions, send it over the server on a cloud, and decode it using ML algorithms and act accordingly.