UNIT-4

Machine Learning

- Machine Learning in the bigger picture
- Areas of machine learning and grades for supervision
- Supervised Learning strategies - regression versus classification
- Unsupervised problem solving-clustering
- Types of Machine Learning:
  - Supervised, Unsupervised
  - Semi-Supervised Learning
  - Reinforcement Learning
- How Supervised Learning works.
- Why the model works on new data.
- Case Study: Recommendation Based Systems, At Microsoft, AI is a Big, Big Deal.

Unit –4 Machine Learning

4.1 Machine Learning:

Artificial Intelligence (AI) is a rapidly evolving technology, made possible by the Internet, which can have a profound impact on our daily lives. AI traditionally refers to the artificial creation of human-like intelligence that can read, discuss, organize, understand, or use natural language. These qualities allow AI to bring greater economic opportunities, while also posing ethical and economic challenges. Since AI is an Internet-enabled technology, the Internet Society recognizes that understanding the opportunities and challenges associated with AI is essential to building an Internet society that people can trust.

Machine learning is a branch of AI. As machine learning is frequently used in products and services, there are other important issues related to users' reliance on the Internet. A number of issues need to be considered when addressing AI, including, social and economic impacts; issues of transparency, fairness, and accountability; new data usage, safety and security considerations, ethical issues; and, and how AI contributes to the creation of new environments.

Algorithms are a sequence of instructions used to solve a problem. Algorithms, developed by computer programmers for new technologies, are the building blocks of the high-quality digital world we see today. Computer algorithms organize large amounts of data into information and services, based on specific commands and rules. It is an important concept to understand, because in machine learning, learning algorithms - not computer programmers - make rules.

Instead of setting up a computer all the time, this method provides computer commands that allow you to read from data without new step-by-step program instructions. This means that computers can be used for new, more complex tasks that cannot be done by hand. Things like photo recognition apps for the visually impaired, or translating photos into speech.

In terms of machines, we can say, in general, that the machine learns whenever it changes its composition, system, or data (based on its input or responding to external information) in such a way that the future is expected performance is improving. Some of these changes, such as the addition of a record in the data generation, it falls nicely into the province of other sectors and we are it is not well understood to be called reading. However, for example, when the function of the speech recognition machine improves after hearing several samples of human speech, we feel appropriate in that context to say that the machine learned. There are several matches between animal and machine learning. Zoologists and psychologists study learning in animals and humans. There are

many techniques in machine learning come from the efforts of psychologists to make more specific their concepts of animal and human learning through computational models. It seems likely also that the theories and techniques being explored by researchers in machine learning may lighten certain aspects of biological learning.

In terms of machines, over-all that the machine learns whenever it changes its composition, system, or data (based on its input or responding to external information) in such a way that the future is predictable performance is improving. For example when the function of the speech recognition machine improves after trial several samples of human speech, we feel appropriate in that context to say that the machine learned.

Machine learning (ML) is a study of computer algorithms that develop automatically with the help of information. It seems to be part of the artificial intelligence. Machine learning algorithms construct a model based on sample data, known as "training data", in order to make predictions or decisions without explicitly planning to do so. Machine learning algorithms are used in a different applications, such as email filtering and computer viewing, where it is tough or impossible to develop general abilities to perform the required tasks.

The subset of machine learning is strictly related to computer statistics, which focus on computer-generated predictions but not all machine learning is mathematical learning. It is study of the application of mathematics brings methods, theoretical and practical contexts into the field of machine learning. Data mining is a coherent field of study, concentrating on the analysis of experimental data by unsupervised learning. The application to business problems, machine learning is also called forensic analytics.

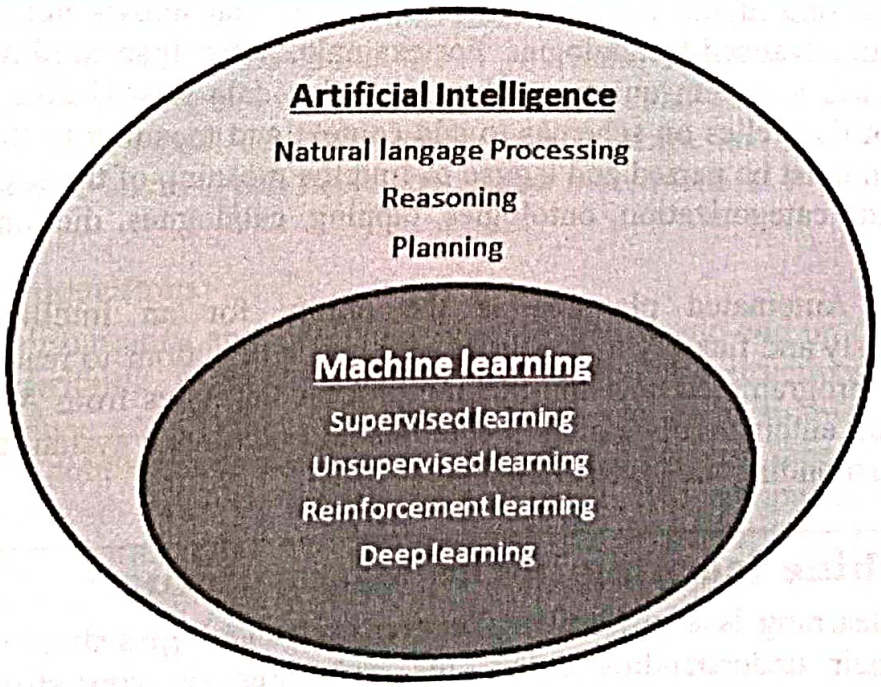

Fig: Relationship between AI and ML

To understand the role of machine learning, we need to give you some perspective. AI, machine learning, and in-depth learning are the terms most often used when talking about big data, math, and advanced technology. AI can be understood as a broader way to describe systems that can "think." For example, thermostats that learn interests or applications that can identify people and what they do in images can be considered as AI programs. As shown in above Figure there are four major AI subsets. In this chapter, we focus on machine learning. However, in order to understand machine learning, it is important to put it in the right way. When testing machine learning, we focus on the ability to read and adapt the model based on data rather than explicit editing.

Reasoning: Machine thinking allows the system to perform assumptions based on data. In fact, consultation helps to fill in the blanks when there are incomplete details. Machine thinking helps to make sense of connected data. For example, if the system has enough data and is asked "What is the safe indoor temperature for eating the drum?" the system will be able to tell you that the answer is 165 degrees. The logic series will be as follows: The edible drum (unlike part of a particular musical instrument) refers to the chicken leg, the chicken leg contains black chicken meat, and the black chicken meat needs to be cooked at 165 degrees, so the response is 165 degrees. Note: In this example, the system was not explicitly trained in the safe internal temperature of the chicken drums. Instead the system used the information that was needed to fill in the data gaps.

Natural Language Processing (NLP): NLP is the ability to train computers to understand both written text and human speech. NLP techniques are needed to capture the meaning of unstructured text from documents or communication from the user. Therefore, NLP is the primary way that systems can interpret text and spoken language. NLP is also one of the fundamental technologies that allows non-technical people to interact with advanced technologies. For example, rather than needing to code, NLP can help users ask a system questions about complex data sets. Unlike structured database information that relies on schemas to add context and meaning to the data, unstructured information must be parsed and tagged to find the meaning of the text. Tools required for NLP include categorization, ontologies, tapping, catalogues, dictionaries, and language models.

Planning: Automated planning is the ability for an intelligent system to act autonomously and flexibly to construct a sequence of actions to reach a final goal. Rather than a pre-programmed decision-making process that goes from A to B to C to reach a final output, automated planning is complex and requires a system to adapt based on the context surrounding the given challenge.

4.2 Machine learning in the bigger picture:

Machine learning is a powerful collection of technologies that can help organizations change their understanding data. This technology is very different from the way companies have traditionally used data. Instead start with a business idea and use it data,

machine learning strategies enable data to create the idea. One of the major advantages of this method is removal business thinking and bias that can lead leaders to agree with a strategy that may not be very effective.

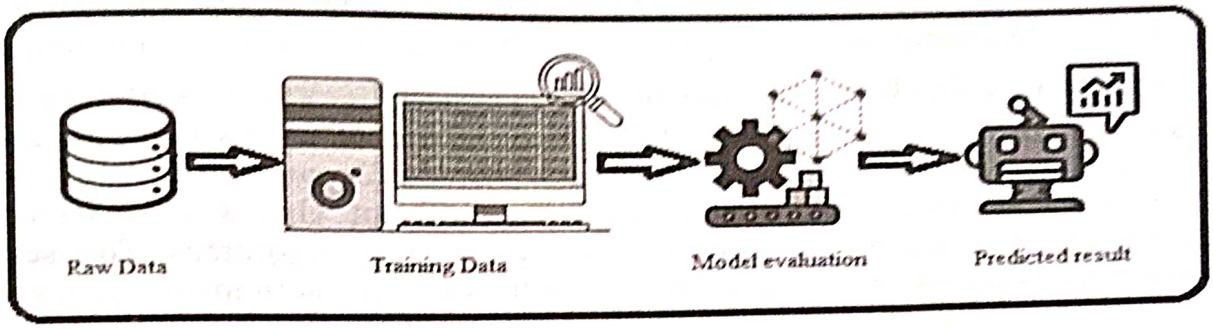

Fig: Workings of Machine learning

Machine learning requires a focus on managing the right data that's well prepared. Organizations should also be able to make choices appropriate algorithms can provide well-designed models. The work does not end there. Machine learning requires a data cycle management, modelling, training, and testing. In this chapter, we focus on technology that supports the machine learning solutions.

➢ **The power of learning the Machine Learning:**

We have made a bold statement that machine learning begins with details and let that data lead you to the idea. How to do business issue a goal? As with all intricate uses development and distribution, requires a planning process understanding the business problem that needs to be solved and collect relevant data sources. How does this approach to planning work affect? In business? When you build sensible apps, you assume that business processes will remain consistent. However, the fact that processes are changing. If you can start with model data, will lead you to systemic and psychological changes. Therefore, machine learning can make application more plentiful it is very powerful and efficient.

➢ **Functions of algorithms:**

There were no discussions about machine learning that would end up outside a category dedicated to algorithms. Algorithms are a set of computer instructions on how to do it collaborate, manage, and modify data. The algorithm can be as simple as the process of adding a number column or as complex as pointing to a person's face in a photo. For the algorithm to work, it must be written as a program that computers can understand. Machine learning algorithms are usually written in one language: Java, Python, or R. One of these languages involves machine learning libraries that support a variety of machine learning skills. In addition, these languages have active user communities constantly coding and discussing ideas, challenges, and methods of business problems. Machine learning algorithms are different from other algorithms. With most algorithms, the

program builder starts by installing algorithm. However, with machine learning the process is investigate. With machine learning, the data itself creates a model. The more data you add to the algorithm, the more complex it becomes the algorithm becomes. As a machine learning algorithm is displayed in additional data, it is able to create more and more intuitive algorithm.

➢ **Types of machine learning algorithms**

Choosing the right algorithm is part of science and part of art. Two data scientists tasked with solving the same business challenge can select different algorithms to approach the same problem. However, to understand the different classes of machine learning algorithms help data scientists identify the best types of algorithms. This section gives you a brief overview of the main type's machine learning algorithms.

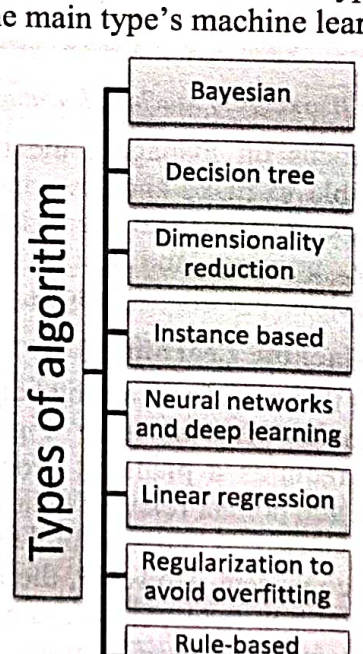

Fig: Types of Algorithms

➢ **Bayesian:**

Bayesian algorithms allow data scientists to record earlier beliefs as to which models should look like, they are not independent of that data he says. With too much focus on model defining data, you might he wonders why people might be interested in Bayesian algorithms. These algorithms are especially useful if you do not have much data values to train confidently model. The Bayesian algorithm will make sense, for example, if you have one previous information on a particular part of the model can therefore encode directly. Let's take the issue of medical imaging diagnosis a system that monitors lung disorders. When a magazine study is published balances the risk of various lung problems based on lifestyle, those possibilities can be modelled.

➢ **Clustering:**

Gathering together is an easy way to understand - objects with the same parameters are grouped together (in groups). All items in the collection are very similar items in other collections. Clustering of the unattended type read because the data has no label. The algorithm translates the parameters form each object and then combine it accordingly.

➢ **Decision tree:**

Decision tree algorithms use branch structure to illustrate the consequences of the decision. Decision trees can be used to draw possible drawings the consequences of the decision. Each node of the decision tree is represented potential outcome. Percentage given to nodes based on chances of an outcome occurring. Decision trees are sometimes used for advertising campaigns. You may want to predict the effect of sending customers again expects a 50% coupon. You can split customers into four parts:

- Motivators who will buy when they gain access.
- Guaranteed items to buy anyway.
- Lost causes that can never be bought.
- Soft customers may react negatively to outreach effort.

If you are posting a marketing campaign, you obviously want to avoid it to send items to three groups because you will not answer, buy anyway, or actually respond negatively. Directing believers will give you the best return on investment (ROI). The decision tree will help you to draw maps of these four groups and set expectations with customers depending on who will respond excellent for marketing campaigns.

➢ **Dimensionality reduction:**

Reducing size helps systems delete incorrect data is useful for analysis. This group of algorithms is used for deletion unwanted data, outliers, and other unusable data. Size reduction can be helpful when analyzing data from sensors and other Internet of Things (IoT) use cases. In IoT systems, there it could be thousands of data points that just tell you that sensor is on. Storing and analyzing that "on" data is useless and will take up significant storage space. Moreover, by removing this unwanted data, machine learning performance the system will improve. Eventually, reducing size will also do help analysts visualize data.

➢ **Instance based:**

Instance-based algorithms are used when you want to categorize new data points are based on the similarity of the training data. This is set algorithms are sometimes called lazy students because there is no training phase. Instead, algorithms are based on the model simply associate new data with training data and separate new ones data points are based on similarity to training data. Occasional support is not well suited to the data sets it contains random variables, invalid data, or data with missing values. Periodically based algorithms can be very helpful in identifying a pattern. For example, learning for example

is applied to chemicals as well biological structural analysis and spatial analysis. Analysis in fields of biology, medicine, chemistry and engineering usually uses various algorithms based on the model.

- **Neural networks and deep learning:**

The neural network attempts to mimic the way the human brain reacts to problems and utilizes layers of interconnected units to read and provide relationships based on visual information. The neural network can have several layers connected. When there is more than one layer hidden in the neural network, it is sometimes called deep learning. Neural network types are able to adapt and learn as data changes. Neural networks are often used when data is unlabelled or unorganized. One of the most important aspects of the use of neural networks is computer vision. In-depth learning is linked today in a variety of forms. Self-driving cars use in-depth learning to help the car understand the environment around the car. As cameras take pictures of natural surroundings, in-depth learning algorithms translate random data to help the system make decisions that are closer to real-time. Similarly, in-depth study is included in applications used by radiologists to help interpret medical imaging.

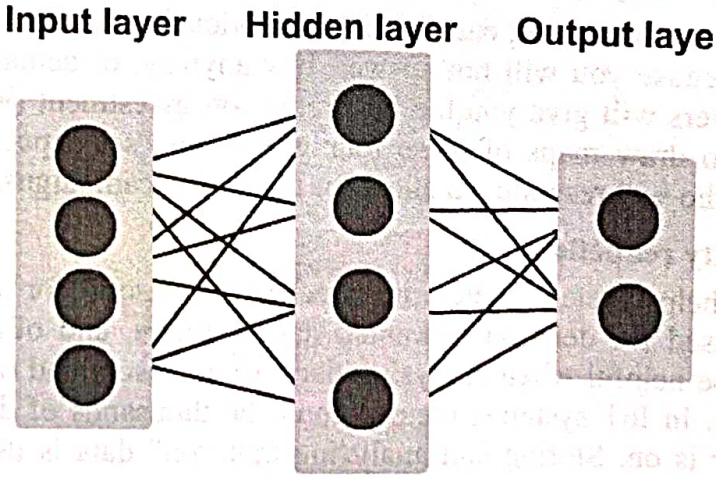

Figure: shows the formation of a neural network. Each layer of neural network filters and change details before passing it will be in the next layer.

- **Linear regression:**

Regression algorithms are widely used in mathematical analysis and are important algorithms used in machine learning. Postponement algorithms help analysts to show relationships between data points. Regression algorithms can measure the strength of a combination between variables in the data set. In addition, regression analysis can be useful in predicting future data rates based on history prices. However, it is important to remember the retreat analysis assumes that the merger is related to contention. Apart

from- to understand the context around the data, regression analysis is possible lead you to inaccurate predictions.

- **Regularization to avoid over-fitting:**

Redesign is a way to change models to avoid a problem excessive skipping. You can add custom to any machine learning model. For example, you can postpone a decision tree model. The redesign makes the models extremely difficult to they tend to be extreme. If the model does not exceed, it will give the wrong one predictions when new data sets are displayed. Excess occurs when the model is designed for specific data set but will have less generalized predictive power data set.

- **Rule-based machine learning:**

Law-based learning algorithms use the rules of relationships to describe data. The legal system can be compared to machine learning programs that form a possible model often applied to all incoming data. In abstract, it is based on rules the systems are easy to understand: When entering X data, do Y. However, as programs begin to be implemented, it is based on rules the machine learning process can be very complicated. For example, the system could include 100 pre-defined rules. As the system meets additional data and is trained, it is possible that there may be hundreds of exemptions from the laws. Icon it is important to be careful when creating a code based on those rules it does not become so complex that it loses its appearance. Imagine how difficult it would be to create a law-abiding law algorithm to use tax code.

- **Process of machine learning systems:**

Through the repeated process of developing and refining the model, selecting the appropriate algorithm, training, and system testing start. Training is a critical step in the process of machine learning. When you train a machine learning program, you know input (eg. customer income, purchase history, location, etc.), and you know your desired goal (predicting the customer output tendency). However, the lesser known is the mathematical operations to convert that raw data into a customer prediction of churn. As the learning algorithm is presented to get more and more customer data, the system will be bigger accurate in predicting customer potential. Training a machine learning algorithm to create an accurate the model can be divided into three steps:

1. Representation:

The algorithm creates a model for converting embedded data to the results you want. As the learning algorithm is presented for more information, it will begin to study the relationship between raw data and what data points are powerful predictors of the result you want.

2. Testing:

As the algorithm creates multiple models, it can be human or the algorithm will need to test and evaluate the models depending on the model that produces the most accurate

predictions. It is important to remember that behind the model is active, will be disclosed in unknown details. Like result, make sure the model is made differently and does not overwork your training data.

3. Development:

After the algorithm created and acquired many models, choose the most efficient algorithm. As you expose the file algorithm for various sets of input data, select the most standard model. The most important part of the training process is getting enough data to be in a position to test your model. Usually the first one passing in training provides mixed results. This means you may need to refine your model or provide more data.

➢ **Machine learning cycle:**

Creating a machine learning program or using a machine learning algorithm is a repetitive process. You can't simply train the model once and leave it alone - data changes, changing preferences, and competitors will appear. Therefore, you need to keep your model fresh when it comes to production. While you will not need to do the same level of training that you needed when you were building the model, you would not think it would be independent. The machine learning cycle continues, and choosing the right machine learning algorithm is one of the steps. The steps of the machine learning cycle are as follows

- Get details: Identifying the right sources of data is the first step in the cycle. In addition, as you develop your machine learning algorithm, consider expanding the targeted data to improve the system.
- Organize data: Make sure your data is clean, secure and controlled. If you perform a wrong machine learning program based on incorrect data, the application will fail.
- Choose a machine learning algorithm: You can have many machine learning algorithms that work on your data and business challenge.
- Train: You need to train the algorithm to create a model. Depending on the type of data and algorithm, the training process can be monitored, supervised or taught reinforcement.
- Analyze: Rate your models to find the most effective algorithm.
- Use: Machine learning algorithms create models that can be used in both cloud and local applications.
- Predictability: After deployment, start making predictions based on new incoming data.
- Check predictions: Check the accuracy of your predictions. The information you collect from analyzing the accuracy of the predictions and then returned to the machine learning cycle to help improve accuracy.

4.3 Areas of Machine Learning and grades for supervision:

Machine learning is the name of modern technology, and it is growing very fast day by day. We use machine learning in our daily lives whether we know it like Google Maps, Google Assistant, Alexa, etc.

Areas of Machine Learning:

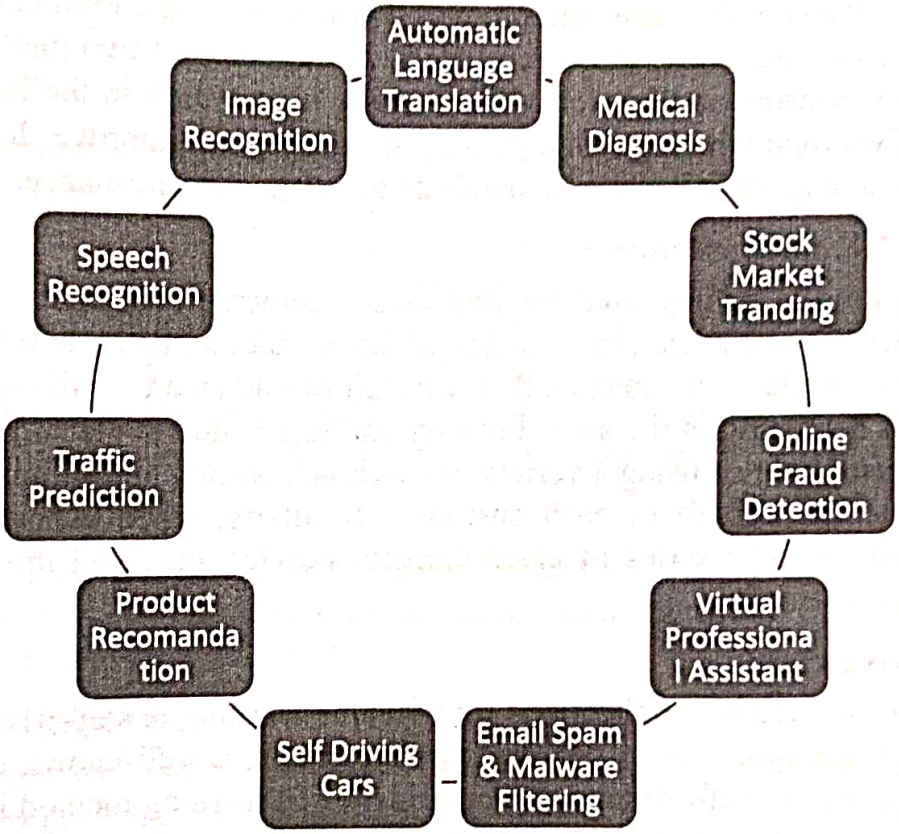

Fig: Application Areas of Machine learning

1. Image recognition:

Image recognition is one of the most common applications for machine learning. Used to identify objects, people, places, digital images, etc. Popular usage case for photo recognition and face detection, auto tagging friend suggestions. Facebook gives us a suggestion feature to tag a friend automatically. Whenever we upload a photo with our Facebook friends, then we automatically get the name tag suggestion, and the technology that supports this is to detect the machine's face detection and algorithm detection. Based on a Facebook project called "Deep Face," which deals with face recognition and photo identification.

2. Speech Recognition:

While using Google, we get the "Voice Search" option, which is subject to speech recognition, and is a popular machine learning program. Speech recognition is the process

of converting voice commands into text, also known as "Speech to text", or "Computer speech recognition." Currently, machine learning algorithms are widely used by various speech recognition systems. Google Assistant Siri, Cortana and Alexa use speech recognition technology to follow voice commands.

3. Traffic Forecast:

If we want to visit a new place, we take the help of Google Maps, which shows us the best route and the shortest route and predicts traffic conditions. Predicts traffic conditions such as traffic cleared, slow, or overcrowded with the help of two methods. Google Map Auto Location Sensors and Sensors. The average time it took in the last few days at the same time. Everyone who uses Google Map helps this app improve. It takes information from the user and sends it back to its database to improve performance.

4. Product Recommendations:

Machine learning is widely used by various e-commerce and entertainment companies such as Amazon, Netflix, etc., to recommend the product to the user. Whenever we search for a particular product on Amazon, then we start to find an ad for the same product while the internet is searching in the same browser and this is due to machine learning. Google understands user interest using a variety of machine learning algorithms and recommends the product to the benefit of each customer. Similarly, when we use Netflix, we get recommendations for a series of entertainment, movies, etc., and this is done with the help of machine learning.

5. Self-driving cars:

One of the most exciting applications for machine learning is self-driving cars. Machine learning plays a major role in self-driving cars. Tesla, a well-known car manufacturing company, operates a self-driving car. An unidentified learning method is used to train car models to find people and objects while driving.

6. Email filtering with Malware:

Whenever we receive a new email, it is automatically filtered as important, normal, and spam. We regularly receive important mail in our inbox with important emails and spam markers in our inbox, and the technology that supports this is machine learning. Below are the spam filters used by Gmail:

- Content Filter
- Header filter
- Normal list list filter
- Law-based filters
- Permission filters

Other machine learning algorithms such as Multi-Layer Perceptron, Decision Tree, and Naïve Bayes edits are used to filter spam email and malware detection.

7. Virtual Personal Assistant:

We have various personal assistants such as Google Assistant, Alexa, Cortana, Siri. As the name suggests, they help us to get information by using our voice commands. These assistants can help us in a variety of ways with our voice commands such as Play music, call someone, Open email, Schedule appointments, etc. These visual aids use machine learning skills as an integral part. These assistants record our voice commands, send them over the server to the cloud, then select them using the ML algorithm and perform accordingly.

8. Internet Fraud Detection:

Machine learning makes our online transactions safer and more secure from fraudulent transactions. Whenever we make an online transaction, there may be a variety of ways in which fraudulent transactions can be made, such as fake accounts, fraudulent ids, and money laundering. So to find out, the Feed Forward Neural network helps us by looking at whether it is a real transaction or a fraudulent transaction. For each real sale, the output is converted into specific hash values, and these values become inputs for the next cycle. For every real sale, there is a pattern that finds a change in the fraudulent material which is why, it finds and makes our online transactions more secure.

9. Stock Market Trading:

Machine learning is widely used in stock market trading. In the stock market, there is always a risk of rising and falling stocks, so in this learning curve a short-term network is used to predict stock market trends.

10. Medical Diagnosis:

In medical science, machine learning is used to diagnose diseases. With this, medical technology is growing rapidly and is able to create 3D models that can predict the exact position of lesions in the brain. It helps in the recovery of brain tissue and other brain-related diseases easily.

11. Automatic language translation:

Nowadays, if we are visiting a new place and we do not know the language then it is not a problem at all, because also this machine learning helps us to translate the text into our known languages. Google's GNMT (Google Neural Machine Translation) offers this feature, which is Neural Machine Learning which translates text into our standard language, and calls it the default translation. The technology behind automatic translation is a sequence of learning algorithm, which is used for image recognition and translates text from one language to another.

- **Grades for Supervision:**

With the learning process we need to have certain findings or data it also known as samples or examples in order to explore possible patterns, hidden in our data. These patterns learned are no longer other than the functions or parameters of the decision.

Machine learning algorithms are often classified as supervised or unsupervised and from these two type further derived into semi-supervised.

- Supervised: All data detection is labelled and algorithms learn output predictions from input data.
- Unsupervised: All data detection has no label and algorithms learn the natural composition from input data.
- Semi-supervised: Some databases are labelled but most of them are usually labelled. Therefore, a combination of supervised and indirect methods is often used.

➢ **Supervised algorithms / methods:**

In this family of models, research needs to have a visual database as well as visual labels / classes. For example, viewing can be pictures of animals and labels the name of an animal (ex: cat, dog, goat etc.)

These types read from a labelled database and are used to predict future events. Through the training process, input is a well-known training data set with its corresponding labels, and the learning algorithm generates targeted activity to ultimately make predictions about new unseen visions one can give the model. The model is able to provide new milestones for any new input after adequate training. The learning algorithm can also compare its output with the expected output and detect errors to modify them properly.

Supervised models can be further classified into regression and classification cases:

- Classification: The problem of classification is when the output flexibility is a category e.g. "Disease" / "no disease".
- Regression: The regression problem is when the output variable is a real continuous value e.g. stock price forecast.

➢ **Unsupervised algorithms / methods:**

In this family of models, research needs to have a database with specific ideas without the need to have labels / visual classes. Unsupervised learning research on how systems can do the job of interpreting hidden structure from unlabelled data. The system does not predict correct output, but instead, it explore data and draw output from data sets to define hidden properties from unlabelled data.

Unsupervised models can be further collected into clustering and Association cases.

- Clustering: The problem of clustering where you want to reveal grouping of data collected, such as collecting animals based on other factors / features e.g. number of legs, color.
- Association: A study of association law is where you want to find association rules as people who buy X and tend to buy Y.

➢ **Supervised machine learning algorithms / methods:**

This family is among the supervised and unsupervised study families. Slightly monitored models use labelled and non-labelled data for training.

➢ **Reinforcement machine learning algorithm / methods:**

This family of models has algorithms that use limited errors such as prizes or penalties. If the error is large, then the penalty is higher and the reward is lower. If the error is small, then the penalty is low and the reward is high. Trial error search and delayed reward are the most appropriate indicators of reinforcement learning. This family of models allows for the automatic determination of good behaviour within a specific context in order to maximize the desired performance.

A reward response is required for the model to learn which action is best and this is known as a reinforcement signal.

4.3.1 Supervised Learning strategies - regression versus classification:

Regression and Classification algorithms are supervised learning algorithms. Both of these algorithms are used to predict machine learning and work with labeled datasets. But the difference between the two is how they are applied to various machine learning problems. The main difference between Regression and Classification algorithms is that Classification algorithm is used to predict the discrete values such as Male or Female, True or False, Spam or Not Spam, etc. whereas Regression algorithm is used to predict continuous values such as price, salary, age, etc. Spam or Not Spam, etc.

- **Classification:**

Classification is the process of finding a function that helps to classify the database into categories based on different parameters. In classification, a computer program is trained in a training database and based on that training, classifies data into different categories. The function of the segmentation algorithm is to find a map function to map input (x) of different output (y).

Example: A good example of understanding the problem of separation is the detection of spam email. The model is trained on the basis of millions of emails in different parameters, and whenever it receives a new email, it indicates whether the email is spam or not. If the email is spam, it is then moved to the spam folder.

Types of ML Classification Algorithms

- ✓ Logistic Regression
- ✓ K-Nearest Neighbours
- ✓ Support Vector Machines
- ✓ Kernel SVM
- ✓ Naïve Bayes
- ✓ Decision Tree Classification

- ✓ Random Forest Classification

- **Regression:**

Regression the process of finding a connection between dependent and independent variables. It helps to predict ongoing trends such as market trend estimates, house price forecasts, etc. The function of the Regression algorithm is to find the map function to map the input variable (x) of the continuous output variable (y).

Example: Suppose we want to predict the weather, so in this case, we will use the Regression algorithm. In weather forecasting, the model is trained in past data, and once the training is completed, it can easily predict future weather.

Types of Regression Algorithm:

- ✓ Simple Linear Regression
- ✓ Multiple Linear Regression
- ✓ Polynomial Regression
- ✓ Support Vector Regression
- ✓ Decision Tree Regression
- ✓ Random Forest Regression

4.3.2 Unsupervised problem solving-clustering:

When faced with real-world problems, most of the time, data will not come with predefined labels, so we will want to develop machine learning models that can better differentiate this data, by discovering for themselves some common features, which will be used to predict classes in new data. The purpose of the merger is to find different groups within the data elements. To do so, integration algorithms detect data formation so that objects of the same group (or group) are more similar than those from different collections.

Many non-invasive problem solving involves collecting objects by looking at the similarity or number of shared features of the visuals, because there is no specific information about the priori categories. This type of strategy is called integration. Apart from these main types of problems, there is a combination of both, called problem-solving problems, in which we can train a set of labeled items and use an index to provide information on non-labeled information during training. Distributing data to anonymous organizations, using three main methods - slide (points close to each other belong to the same category), group (data often form clusters, special slip case).

4.4 Types of machine learning:

Machine learning is a concept that allows the machine to learn from examples and experiences, and that too without being explicitly organized. So instead of coding, what you do is feed the data into a standard algorithm, and the algorithm / machine creates an idea based on the data provided. Let see application of machine learning in our real life:

Ever bought online? So while reviewing a product, are you careful when recommending a product that matches what you want? or note that the person who purchased the product also purchased this product combination. How do they make this recommendation? This is machine learning.

Machine learning is a set of artificial intelligence that focuses on machine learning from their knowledge and making predictions based on their experience. It empowers computers or devices to make data-driven decisions rather than explicitly planning to perform a specific task. These programs or algorithms are designed in a way that learns and develops over time as they are exposed to new data.

Machine Learning algorithm is trained in setting up training data to create a model. When new input details are introduced in the ML algorithm, it makes predictions based on the model. Predictions are tested for accuracy and if accuracy is accepted, the Learning Machine algorithm is included. If accuracy is not acceptable, the Learning Machine algorithm is repeatedly trained with augmented data set. There are many ways to frame this idea, but in particular there are three major categories known:

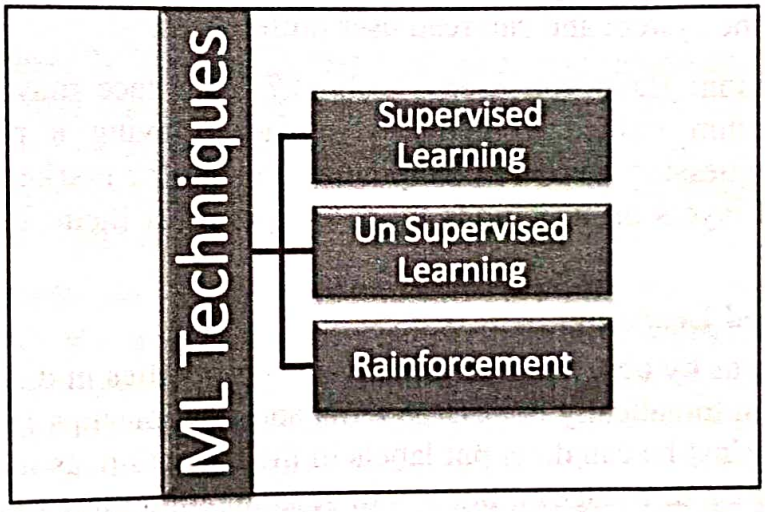

Fig: Types of Machine Learning Techniques

- **Supervised Learning:**

Supervised Learning, where you might think the learning is guided by a teacher. We have a dataset that works as a teacher and its role is to train a model or machine. When a training model can begin to make predictions or decisions when given new data. For example: teaching alphabets to child using flash cards. Which consist of both image and text.

Given the data in the form of labeled examples, we can feed the algorithm for reading these sample label pairs individually, allow the algorithm to predict the label for each sample, and give it feedback on whether it predicts the correct answer or not. Over time, the algorithm will learn to measure the exact nature of the relationship between models

and their labels. When fully trained, the supervised learning algorithm will be able to detect new, unprecedented patterns and predict its own good label.

Supervised learning is often described as work-directed as a result. It focuses on the work of unity, feeding more and more examples in the algorithm until it can perform more accurately in that task. This is the kind of reading you will most likely encounter, as shown in many of the following applications:

Advertising Preferences: Selecting ads that will work best is usually a learning task. Most of the ads you see as you browse the internet are put there because the learning algorithm said they were very popular (and clickable). Additionally, its placement is associated with a specific site or query (if you find yourself using a search engine) mainly because of an educated algorithm that says similarities between ad and placement will work.

Spam Separation: If you are using a modern email program, you may have encountered a spam filter. This spam filter is a curated reading program. Examples of emails and labels (spam / not spam), these programs learn how to pre-filter bad emails so that their user is not harassed by them. Many of these also behave in such a way that the user can provide new labels to the system and can read user preferences.

Face Recognition: Have you used Facebook? Your face may be used in a supervised reading algorithm trained to detect your face. Having a photo-taking, face-to-face program, and guessing who is in the picture (raising the marker) is a supervised process. It has a lot of layers on it, it finds faces and points at them, but it's still being watched anyway.

- **Unsupervised Learning:**

The model learns by observing and discovering properties in data. When a model is given a database, it automatically detects patterns and relationships in the database by creating clusters in it. What he can do is put labels in the collection, as it does not mean that this is a group of apples or mangoes, but it will separate all apples from mangoes. Suppose we present pictures of apples, bananas and mangoes in the model, so what it does, based on other patterns and relationships build clusters and separate the databases from those collections. Now when new data is added to the model, it adds it to one of the created collections.

Unsupervised learning is very different from supervised learning. No labels. Instead, our algorithm will be fed with a lot of detail and provided with tools to understand data structures. From there, it can learn to collect, compile, and / or organize data in such a way that a person (or other intelligent algorithm) can enter and make sense of newly edited data.

For example, what if we had a large database of all the research papers that have already been published and we had surveyed study programs that knew how to collect them in such a way that you were always aware of current trends within a particular research

domain. Now, you start your own research project, connecting your work to this network that you can see the algorithm. As you write your work and take notes, the algorithm makes suggestions for you about related tasks, tasks you would like to mention, and it works that can help you move forward in that research area. With such a tool, your product can be maximized.

Because unsupervised learning is based on data and its properties, we can say that unchecked reading is driven by data. Outcomes from unsupervised learning activity are controlled by data and formatted. Some areas where you can see that unsupervised reading restrictions are:

Recommendation Programs: If you have ever used YouTube or Netflix, you may have encountered a video recommendations program. These applications are usually installed on an unmanaged domain. We know things about videos, maybe their length, genre, etc. We also know the watch history of multiple users. Considering users who have watched similar videos like you and enjoyed other videos you have not yet seen, the recommendation system can detect this data relationship and provide you with such a suggestion.

Purchasing Practices: It is possible that your purchase practices are contained in the database elsewhere and that the data is actively purchased and sold at this time. These purchase practices can be applied to learning algorithms that can be controlled to gather customers in the same shopping categories. This helps companies market in these aggregated categories and can serve as promotional programs.

User Logging: A little user experience, but still very effective, we can use unchecked reading to collect logs and user problems. This can help companies identify key themes in the problems their customers face and address these issues, by improving product or designing FAQ to address common issues. In any case, it is a work in progress and if you have ever caused a problem with a product or submitted a bug report, it may have been fed an unattended reading algorithm to integrate it with other similar issues.

Reinforcement:

It is the agent's ability to interact with nature and find out what the best result is. It follows the theory of hit and trial method. The agent is rewarded for a point with a correct or incorrect answer, and on the basis of good reward points earn the model trains themselves and once trained it is ready to predict the new data that are being introduced to it.

Reinforce learning is very diverse compared to supervised and supervised learning. Where we can easily see the relationship between supervised and unsupervised (presence or absence of labels), relationships in strengthening teaching are less positive. Some people try to tie the knot of reading close to these by describing them as a type of reading that relies on time-based label sequences, however, my view is that that just makes things very confusing.

I prefer to look at strengthening reinforcement as learning from mistakes. Fit the reinforcement learning algorithm in any environment and it will make many mistakes in the beginning. As long as we provide a specific type of signal in an algorithm that combines positive behaviour with positive signals as well as negative and negative behaviours, we can strengthen our algorithm to select positive behaviours rather than negative ones. Over time, our learning algorithm learns to make smaller mistakes than before.

The strengthening of learning is largely driven by morality. It has influences from the field of neuroscience and psychology. However, to really understand the reinforcement of reading, let's separate the concrete example. Let's take a look at teaching an agent to play the game Mario.

For any reinforcement learning problem, we need an agent and a location and a way to connect the two items using a feedback loop. To connect the agent to the environment, we give you a set of actions that you can take that affect the environment. To connect the environment to the agent, we are able to continuously extract two signals from the agent: an update state and a reward (our signal to strengthen the behaviour). In Mario's game, our agent is our learning algorithm and our environment is a game. Our agent has a set of actions. These will be our buttons. Our updated status will be for each frame of the game as time goes on and our reward signal will be a change in points. As long as we connect all of these things together, we will have established a strong sense of learning to play the game Mario.

- **Where is reinforcement learning in the real world?**

Video Games: One of the most common areas to look at is strengthening learning to play games. Check out the enhancement learning program Google, AlphaZero and AlphaGo has learned to play Go game. Our example of Mario is a classic example.

Industrial Simulation: In many robotic systems (think assembly lines), it helps our machines learn to complete their tasks without having to go through complex procedures. This can be a cheap and safe option; it can even be a habit of failure. We can also encourage our machines to use less electricity, to save money. In addition, we can start all of this in imitation so that we do not waste money in case we break our machine.

Resource Management: Strengthening learning is good for traveling in complex environments. It can handle the need to balance certain needs. Take, for example, Google data centers. They have used the strengthening of learning to balance the need to meet our energy needs, but they have done it as well as possible, reducing significant costs. How does this affect us and the average person? Cheap storage costs for our data also have a small impact on the environment we all share.

4.5 Why the model works on new data.

Thanks to new computer technologies, machine learning today is not the same as machine learning in the past. Born of pattern recognition and the idea that computers can learn without being programmed to perform certain tasks researchers interested in artificial intelligence wanted to see if computers could learn from data. The iterative aspect of machine learning is important because as models are presented in new data, they are able to adapt independently. They learn from previous statistics to produce reliable, repetitive decisions and results. It is a science that is not new - but one that has gained new momentum.

While many electronic learning algorithms have long existed, the ability to automate complex mathematical calculations into big data - often, quickly and quickly - is a recent development. Here are a few widely distributed examples of machine learning applications you may know:

- Google's fast-paced, self-driving car? Total machine learning.
- Online recommendations like those from Amazon and Netflix? Applications for machine learning in everyday life.
- Do you know what customers are saying about you on Twitter? Machine learning is integrated into the development of language rules.
- Fraud detection? One of the most obvious, practical things in our world today.

4.6 Case Study: Recommendation Based Systems, At Microsoft, AI is a Big, Big Deal.

For better or for worse, depending on your point of view in terms of technological impact on work, artificial intelligence has long been seen as a technology to help companies do more with fewer people. That is the key to making companies 'shine,' as we refer to in this issue of TCS Perspectives to run and grow a business without rising geometric salaries.

However, Joseph Sirosh, vice president of Microsoft Corp. In charge of information management and machine learning, he believes that AI - and most importantly, machine learning - will be needed to help companies run business processes where there are not enough people to make it a priority. Many of those processes handle large and continuous volumes of digital data.

A good example of this is how big companies protect their computer systems from attackers, hackers who try to infiltrate their networks, malware that infiltrates their email systems and web browsers, and more. "All of those things today are very well received in real time and automatically, using machine learning skills," he explains. "It is a fact that the modern algorithms for learning computer technology are what keep Microsoft's cyber infrastructure information secure. There are not many people [at Microsoft] roaming around trying to find out if something is wrong. It just wouldn't go up. "

Without the automatic form of machine learning to detect cyber-attacks, Microsoft and other major companies would have difficulty defending themselves quickly before the damage was done, he notes. "Without this kind of automated system, it would be very difficult for these things to be found very quickly to prevent them," he said. When people open up a virus in a certain part of the world that starts to spread on networks and infect PCs, if you look at the data from these machines, you can understand what is happening - if there is an algorithm that goes on in secret we warn you of big changes taking place.

Monitoring its computer networks is not the only place Microsoft has been using its ever-expanding and advanced technology. Sirosh and others have been employed by Microsoft for several years to inject machine learning models into Microsoft products and services. Products such as the Bing search engine and Microsoft's entry into the digital assistant market (Cortana) are full of machine learning capabilities to help the computer system become 'smarter' itself, without the need for human editors.

Take Bing, a search engine in 2009 that was the third largest in the market on Google and Yahoo, according to market tracker ComScore. Back then, Microsoft introduced a data scientist (Qi Lu, now Microsoft EVP) who recommended the company's search engineers to develop machine learning algorithms that would automatically and continuously improve Bing's ability to call relevant content. It has also suggested that Microsoft create a database that stores all its search data, a critical piece of machine learning.

Bing engineers follow his advice, and great things happen. By producing more relevant search results for Bing users, between 2009 and 2015 Bing's market share doubled to 20%. In addition, Microsoft's search business has grown to more than $ 1 billion a quarter and has become more profitable.

"The quality of the quality results produced by the Bing search engine depends entirely on the machine learning models behind it," explains Sirosh. "Machine learning tests search questions and what people click on. Then they build a very powerful model and distribute it in a few details that are designed to ask very quickly. "

The result is that all the search results you get by typing words into Bing are available and calculated by machine learning model. "They do a great deal of quality improvement for our search customers," Sirosh said. "This is just one example where machine learning is built entirely on product design and has become one of its major dividers."

This is one of the many ways in which Microsoft has incorporated machine learning into its technological products and services. It contains mid-range strategies by Microsoft CEO Satya Nadella to continue growing the 41-year-old company far from personal computer. As a Bloomberg Businessweek article put it earlier this year, Nadella "has been sprinkling machine learning like fairy dust on everything his company touches."1 Sirosh and many others at Microsoft are there to make that happen.

1. Dina Bass, "Inside Microsoft, Where Lie Detection Is a Killer App," Bloomberg Businessweek, Feb. 22, 2016. Accessed June 24, 2016. http://www.bloomberg.com/news/articles/2016-02-22/inside-the-new-microsoft-where-lie-detection-is-a-killer-app

Exercises

❖ Answer the following Questions in brief.

1. What is machine learning? How it is relates to AI?
2. Explain in detail the workings of machine learning.
3. Explain the types of machine learning algorithms.
4. Write the process and cycle of the machine learning system.
5. Discuss the types of machine learning techniques.
6. Explain the areas of machine learning.
7. Discuss about the supervised and unsupervised algorithm.
8. Differentiate the supervised and unsupervised techniques of machine learning.
9. Differentiate the regression vs. classification model.
10. Why model works on new data?

❖ Multipal choice Questions - MCQs:

1. Application of machine learning methods to large databases is called -
   A. big data analysis B. artificial intelligence
   C. data mining D. internet of things
2. If machine learning model output involves target variable then that model is called as
   A. descriptive model B. predictive model
   C. reinforcement learning D. all of the above
3. In what type of learning labelled training data is used
   A. unsupervised learning B. active learning
   C. reinforcement learning D. supervised learning
4. Which of the following is the best machine learning method?
   A. scalable B. accuracy
   C. fast D. all of the above
5. What characterize unlabeled examples in machine learning
   A. there is no prior knowledge B. there is no confusing knowledge

C. there is plenty of confusing knowledge D. there is prior knowledge

6. Data used to build a data mining model.

A. test data B. validation data

C. training data D. hidden data

7. The problem of finding hidden structure in unlabeled data is called...

A. unsupervised learning B. supervised learning

C. reinforcement learning D. none of the above

8. Of the Following Examples, Which would you address using an supervised learning Algorithm?

A. given a set of news articles found on the web, group them into set of articles about the same story.

B. given email labeled as spam or not spam, learn a spam filter

C. given a database of customer data, automatically discover market segments and group customers into different market segments.

D. find the patterns in market basket analysis

9. You are given reviews of few amazon series marked as positive, negative and neutral. Classifying reviews of a new amazon series is an example of

A. semi supervised learning B. unsupervised learning

C. supervised learning D. reinforcement learning

10. Following are the types of supervised learning.

A. classification B. regression

C. subgroup discovery D. all of the above

11. The output of training process in machine learning is

A. dataset B. machine learning algorithm

C. machine learning model D. accuracy

12. Type of matrix decomposition model is

A. predictive model B. descriptive model

C. logical model D. none of the above

13. Machine Learning is the autonomous acquisition of knowledge through the use of computer programs. (True / False)

14. ML is a field of AI consisting of learning algorithms that?

A. Improve their performance B. At executing some task

B. Over time with experience D. All of the above

15. Different learning methods in the ML do not include Introdution. (True / False)

16. Some telecommunication company wants to segment their customers into distinct groups ,this is an example of

A. unsupervised learning B. reinforcement learning

C. supervised learning D. data extraction

17. Which learning Requires Self Assessment to identify patterns within data?

A. supervised learning B. reinforcement learning

C. unsupervised learning D. data extraction

18. In simple term, machine learning is

A. training based on historical data B. prediction to answer a query

C. both a and b?? D. none

19. If machine learning model output does not involves target variable then that model is called as

A. regression model B. predictive model

C. reinforcement learning D. descriptive model

20. Following are the descriptive models

A. clustering B. association rule

C. both a and b D. none

Answers:

1. C	2.B	3.D	4. B	5. C
6. C	7. A	8. B	9. C	10. D
11. C	12. B	13. True	14. D	15. True
16. A	17. C	18. C	19. D	20. C

GUJARAT UNIVERSITY

April – 2021

B.C.A SEM – VI

CC- 309 Introduction to Artificial Intelligence and Machine Learning (New Course)

Time: 2:00 Hrs. **Model Paper** **Total Marks: 50**

Instruction: All Questions of **section I** carry equal marks.

Attempt any **two** Questions in **section I**

Question 5 in section II **is COMPULSORY,** Attempt any **Five.**

SECTION - I

Q. 1 A: Write all the eight, different definitions of Artificial Intelligence. 10

B: Short note on State of the Art applications of Artificial Intelligence. 10

Q. 2 A: 1. Explain the difference between BFS and DFS? 10
2. Explain Informed and Uninformed strategies with a Toy problem.

B: What is the concept of Greedy BFS and its algorithm? Explain with real time example? 10

Q. 3 A: Define language model. Define types of language model. 10

B: What are unigrams, bigrams, trigrams, and n-grams in NLP? 10

Q. 4 A: Explain the types of machine learning algorithms. 10

B: Write the process and cycle of the machine learning system. 10

SECTION- II

Q.5 Answers the following Questions (Any 5 x 2 marks each). 10

1. Select the most appropriate situation for that a blind search can be used.

A. Real-life situation B. Small Search Space

C. Complex game D. All of the above

2. If a robot is able to change its own trajectory as per the external conditions, then the robot is considered as the _____.

A. Mobile B. Non-Servo C. Open Loop D. Intelligent

3. Greedy search strategy chooses the node for expansion in _____

A. Shallowest B. Deepest

C. The one closest to the goal node D. Minimum heuristic cost

4. What is Initial state + Goal state in Search Terminology?

A. Problem Space B. Problem Instance

C. Problem Space Graph D. Admissibility

5. Machine Learning is the autonomous acquisition of knowledge through the use of computer programs. (True / False)

6. What is full form of NLP?

A. Nature Language Understanding B. Natural Long Processed

C. Natural Language Processing D. None of the Above

7. What are the input and output of an NLP system?

A. Speech and noise B. Speech and Written Text

C. Noise and Written Text D. Noise and value

8. What is compression?

A. To compress something by pressing it very hardly

B. To minimize the time taken for a file to be downloaded

C. To reduce the size of data to save space

D. To convert one file to another Answer

9. Different learning methods in the ML do not include Introduction. (True / False)

10. Which of the following is the best machine learning method?

A. scalable B. accuracy

C. fast D. all of the above