BCA
Semester-6

USEFUL FOR COMPUTER WORLD BCA STUDENT

Introduction to

Artificial Intelligence and Machine Learning

Subject Code : CC - 309

Dr. Manali Brahmbhatt
Ph.D. (Computer Science)
St. Xavier's College, Navrangpura, A'bad.

Prof. Deepal Shah
MCA (Master in Computer Application), OCP9i-DBA.
JG-CCA, ASIA-Campus,
Opp. Drive-In-Cinema, Thaltej -Ahmedabad 54.

Prof. Mital Shingala
MCA (Master in Computer Application)
Navgujarat College of Computer Applications, Gandhinagar.

Prof. Himangni Rathore
B.E. (IT), M.Tech (Software Systems)
Regional Commissioner of Municipalities, Gandhinagar.
Former Assistant Professor,
Navgujarat College of Computer Applications, Gandhinagar.

COMPUTER WORLD PUBLICATION
A Division of Live Education System Pvt. Ltd.

COMPUTER WORLD
43, 5th Floor, SANIDHYA Complex, Nr. M. J. Library,
Opp. Sanyas Ashram, Ashram Road, Ahmedabad-09.
Mobile : 9725019114, 9725022917, 9725020595, 9825020595
URL : www.computerworld.ind.in | e-Mail :info@computerworld.ind.in

Publish By :

COMPUTER WORLD

A Division of Live Education System Pvt. Ltd.

An ISO 9001 : 2008 Certified Company

Kalpesh Patel - kalpesh@computerworld.ind.in

Prepare by : Computer World Research Department

Dr. Manali Brahmbhatt, Prof. Himangni Rathore,
Prof. Deepal Shah, Prof. Mital Shingala, Ravindrabhai

Price : ₹ 99/-

Book Code : CEBCA124

Edition : 1st

Scanned with CamScanner

Index

Scanned with CamScanner

Scanned with CamScanner

UNIT-1

Fundamentals of Artificial intelligence and Intelligent Agent

- What is AI?
  - Acting humanly: The Turing Test approach
  - Thinking humanly: The cognitive modeling approach
  - Thinking rationally: The "laws of thought" approach
  - Acting rationally: The rational agent approach
- State of Art (Applications of AI)
- Agents and Environments
- The Concept of Rationality
- The Nature of Environment
- The Structure of Agents.
- Case Study: Create a new health care market with AI

Scanned with CamScanner

❖ Introduction:

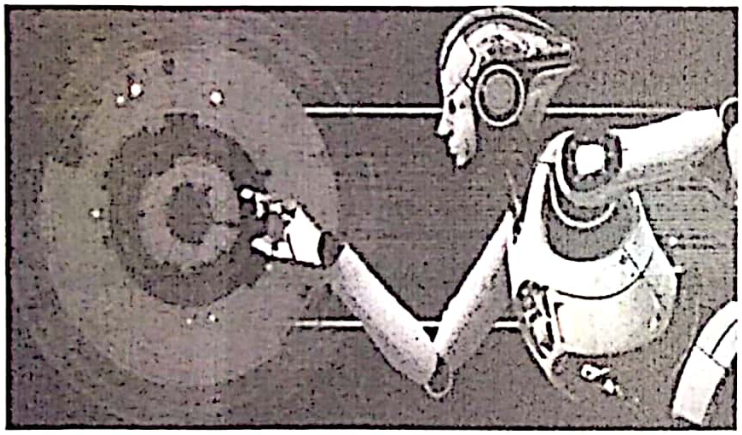
The Artificial Intelligence tutorial provides an introduction to AI which will help you to understand the concepts behind Artificial Intelligence. In this tutorial, we have also discussed various popular topics such as History of AI, applications of AI, deep learning, machine learning, natural language processing, Reinforcement learning, Q-learning, Intelligent agents, Various search algorithms, etc.

Our AI tutorial is prepared from an elementary level so you can easily understand the complete tutorial from basic concepts to the high-level concepts.

1.1 What is Artificial Intelligence?

In today's world, technology is growing very fast, and we are getting in touch with different new technologies day by day.

Here, one of the booming technologies of computer science is Artificial Intelligence which is ready to create a new revolution in the world by making intelligent machines. The Artificial Intelligence is now all around us. It is currently working with a variety of subfields, ranging from general to specific, such as self-driving cars, playing chess, proving theorems, playing music, Painting, etc.



AI is one of the fascinating and universal fields of Computer science which has a great scope in future. AI holds a tendency to cause a machine to work as a human.

Artificial Intelligence is composed of two words **Artificial** and **Intelligence**, where Artificial defines *"man-made,"* and intelligence defines *"thinking power"*, hence AI means *"a man-made thinking power."*

So, we can define AI as:

"It is a branch of computer science by which we can create intelligent machines which can behave like a human, think like humans, and able to make decisions."

Scanned with CamScanner

Artificial Intelligence exists when a machine can have human based skills such as learning, reasoning, and solving problems With Artificial Intelligence you do not need to preprogram a machine to do some work, despite that you can create a machine with programmed algorithms which can work with own intelligence, and that is the awesomeness of AI.

It is believed that AI is not a new technology, and some people says that as per Greek myth, there were Mechanical men in early days which can work and behave like humans.

Definitions of artificial intelligence according to eight recent textbooks are shown in the table below. These definitions vary along two main dimensions. The ones on top are concerned with thought processes and reasoning, whereas the ones on the bottom address behavior. Also, the definitions on the left measure success in terms of human performance, whereas the ones on the right measure against an ideal concept of intelligence, which we will call rationality. A system is rational if it does the right thing.

``The exciting new effort to make computers think ... machines with minds, in the full and literal sense" (Haugeland, 1985) ``The automation of activities that we associate with human thinking, activities such as decision-making, problem solving, learning ..." (Bellman, 1978)	``The study of mental faculties through the use of computational models" (Charniak and McDermott, 1985) ``The study of the computations that make it possible to perceive, reason, and act" (Winston, 1992)
``The art of creating machines that perform functions that require intelligence when performed by people" (Kurzweil, 1990) ``The study of how to make computers do things at which, at the moment, people are better" (Rich and Knight, 1991)	``A field of study that seeks to explain and emulate intelligent behavior in terms of computational processes" (Schalkoff, 1990) ``The branch of computer science that is concerned with the automation of intelligent behavior" (Luger and Stubblefield, 1993)

This gives us four possible goals to pursue in artificial intelligence:

Systems that think like humans.	Systems that think rationally.
Systems that act like humans	Systems that act rationally

Historically, all four approaches have been followed. As one might expect, a tension exists between approaches centered around humans and approaches centered around rationality. (We should point out that by distinguishing between human and rational behavior, we are not suggesting that humans are necessarily ``irrational" in the sense of ``emotionally unstable" or ``insane." One merely need note that we often make mistakes; we are not all

Scanned with CamScanner

chess grandmasters even though we may know all the rules of chess; and unfortunately, not everyone gets an A on the exam. Some systematic errors in human reasoning are cataloged by Kahneman et al..) A human-centered approach must be an empirical science, involving hypothesis and experimental confirmation. A rationalist approach involves a combination of mathematics and engineering. People in each group sometimes cast aspersions on work done in the other groups, but the truth is that each direction has yielded valuable insights. Let us look at each in more detail.

➢ Why Artificial Intelligence?

Before Learning about Artificial Intelligence, we should know that what is the importance of AI and why should we learn it. Following are some main reasons to learn about AI:

- With the help of AI, you can create such software or devices which can solve real-world problems very easily and with accuracy such as health issues, marketing, traffic issues, etc.
- With the help of AI, you can create your personal virtual Assistant, such as Cortana, Google Assistant, Siri, etc.
- With the help of AI, you can build such Robots which can work in an environment where survival of humans can be at risk.
- AI opens a path for other new technologies, new devices, and new Opportunities.

➢ Goals of Artificial Intelligence:

Following are the main goals of Artificial Intelligence:

1. Replicate human intelligence
2. Solve Knowledge-intensive tasks
3. An intelligent connection of perception and action
4. Building a machine which can perform tasks that requires human intelligence such as:
   - Proving a theorem
   - Playing chess
   - Plan some surgical operation
   - Driving a car in traffic
5. Creating some system which can exhibit intelligent behavior, learn new things by itself, demonstrate, explain, and can advise to its user.

➢ What Comprises to Artificial Intelligence?

Artificial Intelligence is not just a part of computer science even it's so vast and requires lots of other factors which can contribute to it. To create the AI first we should know that how intelligence is composed, so the Intelligence is an intangible part of our brain which is a combination of **Reasoning, learning, problem-solving perception, language understanding, etc.**

Scanned with CamScanner

To achieve the above factors for a machine or software Artificial Intelligence requires the following discipline:

- Mathematics
- Biology
- Psychology
- Sociology
- Computer Science
- Neurons Study
- Statistics

➢ Advantages of Artificial Intelligence:

Following are some main advantages of Artificial Intelligence:

- **High Accuracy with less errors:** AI machines or systems are prone to less errors and high accuracy as it takes decisions as per pre-experience or information.
- **High-Speed:** AI systems can be of very high-speed and fast-decision making, because of that AI systems can beat a chess champion in the Chess game.
- **High reliability:** AI machines are highly reliable and can perform the same action multiple times with high accuracy.
- **Useful for risky areas:** AI machines can be helpful in situations such as defusing a bomb, exploring the ocean floor, where to employ a human can be risky.
- **Digital Assistant:** AI can be very useful to provide digital assistant to the users such as AI technology is currently used by various E-commerce websites to show the products as per customer requirement.
- **Useful as a public utility:** AI can be very useful for public utilities such as a self-driving car which can make our journey safer and hassle-free, facial recognition for security purpose, Natural language processing to communicate with the human in human-language, etc.

➢ Disadvantages of Artificial Intelligence:

Every technology has some disadvantages, and thesame goes for Artificial intelligence. Being so advantageous technology still, it has some disadvantages which we need to keep in our mind while creating an AI system. Following are the disadvantages of AI:

- **High Cost:** The hardware and software requirement of AI is very costly as it requires lots of maintenance to meet current world requirements.
- **Can't think out of the box:** Even we are making smarter machines with AI, but still they cannot work out of the box, as the robot will only do that work for which they are trained, or programmed.

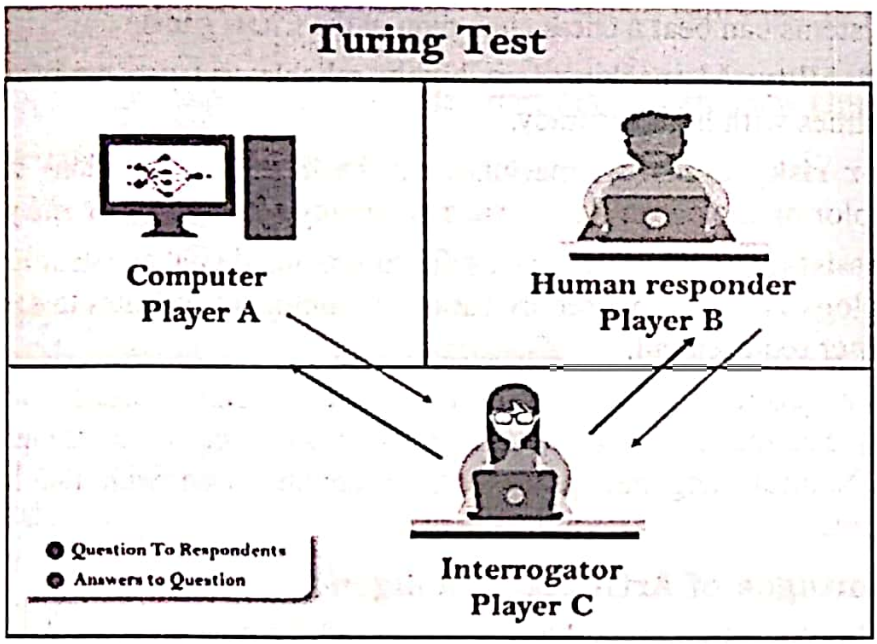
Scanned with CamScanner

- **No feelings and emotions:** AI machines can be an outstanding performer, but still it does not have the feeling so it cannot make any kind of emotional attachment with human, and may sometime be harmful for users if the proper care is not taken.
- **Increase dependency on machines:** With the increment of technology, people are getting more dependent on devices and hence they are losing their mental capabilities.
- **No Original Creativity:** As humans are so creative and can imagine some new ideas but still AI machines cannot beat this power of human intelligence and cannot be creative and imaginative.

➢ Turing Test in AI:

In 1950, Alan Turing introduced a test to check whether a machine can think like a human or not, this test is known as the Turing Test. In this test, Turing proposed that the computer can be said to be an intelligent if it can mimic human response under specific conditions.

Turing Test was introduced by Turing in his 1950 paper, "Computing Machinery and Intelligence," which considered the question, "Can Machine think?"



The Turing test is based on a party game "Imitation game," with some modifications. This game involves three players in which one player is Computer, another player is human responder, and the third player is a human Interrogator, who is isolated from other two players and his job is to find that which player is machine among two of them.

Consider, Player A is a computer, Player B is human, and Player C is an interrogator. Interrogator is aware that one of them is machine, but he needs to identify this on the basis of questions and their responses.

The conversation between all players is via keyboard and screen so the result would not depend on the machine's ability to convert words as speech.The test result does not

Scanned with CamScanner

depend on each correct answer, but only how closely its responses like a human answer. The computer is permitted to do everything possible to force a wrong identification by the interrogator.

The questions and answers can be like:

Interrogator: Are you a computer?

PlayerA (Computer): No

Interrogator: Multiply two large numbers such as (256896489*456725896)

Player A: Long pause and give the wrong answer.

In this game, if an interrogator would not be able to identify which is a machine and which is human, then the computer passes the test successfully, and the machine is said to be intelligent and can think like a human. "In 1991, the New York businessman Hugh Loebner announces the prize competition, offering a $100,000 prize for the first computer to pass the Turing test. However, no AI program to till date, come close to passing an undiluted Turing test".

- **Chatbots to attempt the Turing test:**

ELIZA: ELIZA was a Natural language processing computer program created by Joseph Weizenbaum. It was created to demonstrate the ability of communication between machine and humans. It was one of the first chatterbots, which has attempted the Turing Test.

Parry: Parry was a chatterbot created by Kenneth Colby in 1972. Parry was designed to simulate a person with **Paranoid schizophrenia**(most common chronic mental disorder). Parry was described as "ELIZA with attitude." Parry was tested using a variation of the Turing Test in the early 1970s.

Eugene Goostman: Eugene Goostman was a chatbot developed in Saint Petersburg in 2001. This bot has competed in the various number of Turing Test. In June 2012, at an event, Goostman won the competition promoted as largest-ever Turing test content, in which it has convinced 29% of judges that it was a human.Goostman resembled as a 13-year old virtual boy.

- **The Chinese Room Argument:**

There were many philosophers who really disagreed with the complete concept of Artificial Intelligence. The most famous argument in this list was **"Chinese Room."**

In the year **1980, John Searle** presented **"Chinese Room"** thought experiment, in his paper **"Mind, Brains, and Program,"** which was against the validity of Turing's Test. According to his argument, **"Programming a computer may make it to understand a language, but it will not produce a real understanding of language or consciousness in a computer."**

Scanned with CamScanner

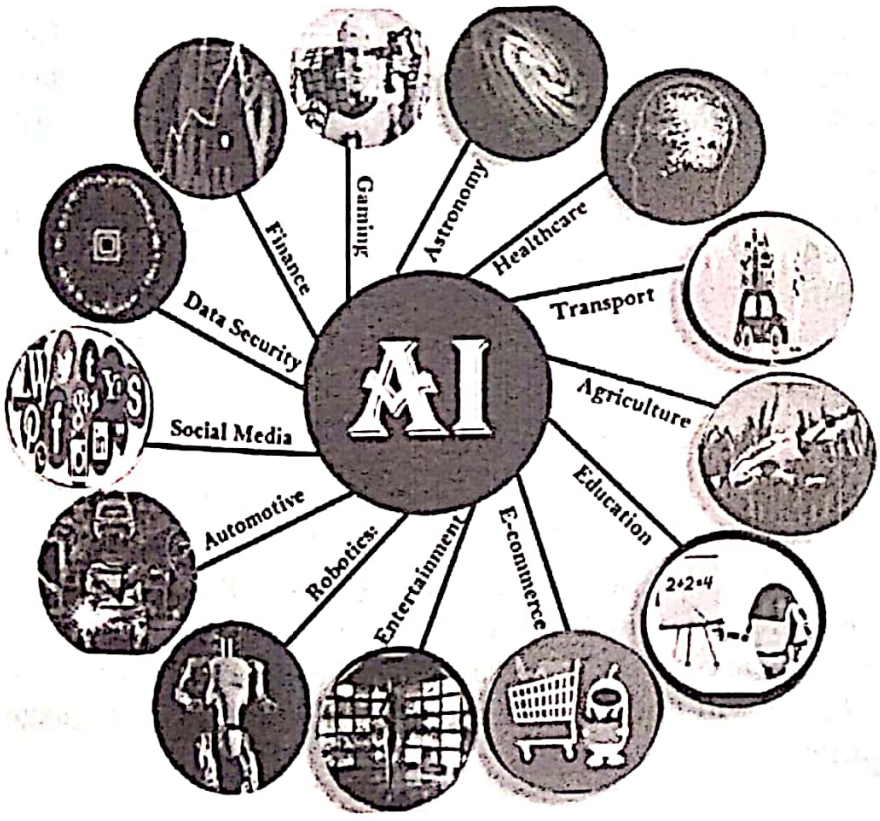
He argued that Machine such as ELIZA and Parry could easily pass the Turing test by manipulating keywords and symbol, but they had no real understanding of language. So it cannot be described as "thinking" capability of a machine such as a human.

- **Features required for a machine to pass the Turing test:**
  - **Natural language processing:** NLP is required to communicate with Interrogator in general human language like English.
  - **Knowledge representation:** To store and retrieve information during the test.
  - **Automated reasoning:** To use the previously stored information for answering the questions.
  - **Machine learning:** To adapt new changes and can detect generalized patterns.
  - **Vision (For total Turing test):** To recognize the interrogator actions and other objects during a test.
  - **Motor Control (For total Turing test):** To act upon objects if requested.

1.2 State of Art (Application of AI):

Artificial Intelligence has various applications in today's society. It is becoming essential for today's time because it can solve complex problems with an efficient way in multiple industries, such as Healthcare, entertainment, finance, education, etc. AI is making our daily life more comfortable and fast.

Following are some sectors which have the application of Artificial Intelligence:



Scanned with CamScanner

1. AI in Astronomy:

- Artificial Intelligence can be very useful to solve complex universe problems. AI technology can be helpful for understanding the universe such as how it works, origin, etc.

2. AI in Healthcare:

- In the last, five to ten years, AI becoming more advantageous for the healthcare industry and going to have a significant impact on this industry.
- Healthcare Industries are applying AI to make a better and faster diagnosis than humans. AI can help doctors with diagnoses and can inform when patients are worsening so that medical help can reach to the patient before hospitalization.

3. AI in Gaming:

- AI can be used for gaming purpose. The AI machines can play strategic games like chess, where the machine needs to think of a large number of possible places.

4. AI in Finance:

- AI and finance industries are the best matches for each other. The finance industry is implementing automation, chatbot, adaptive intelligence, algorithm trading, and machine learning into financial processes.

5. AI in Data Security:

- The security of data is crucial for every company and cyber-attacks are growing very rapidly in the digital world. AI can be used to make your data more safe and secure. Some examples such as AEG bot, AI2 Platform,are used to determine software bug and cyber-attacks in a better way.

6. AI in Social Media:

- Social Media sites such as Facebook, Twitter, and Snapchat contain billions of user profiles, which need to be stored and managed in a very efficient way. AI can organize and manage massive amounts of data. AI can analyze lots of data to identify the latest trends, hashtag, and requirement of different users.

7. AI in Travel & Transport:

- AI is becoming highly demanding for travel industries. AI is capable of doing various travel related works such as from making travel arrangement to suggesting the hotels, flights, and best routes to the customers. Travel industries are using AI-powered chatbots which can make human-like interaction with customers for better and fast response.

8. AI in Automotive Industry:

- Some Automotive industries are using AI to provide virtual assistant to their user for better performance. Such as Tesla has introduced TeslaBot, an intelligent virtual assistant.

Scanned with CamScanner

- Various Industries are currently working for developing self-driven cars which can make your journey more safe and secure.

9. AI in Robotics:

- Artificial Intelligence has a remarkable role in Robotics. Usually, general robots are programmed such that they can perform some repetitive task, but with the help of AI, we can create intelligent robots which can perform tasks with their own experiences without pre-programmed.
- Humanoid Robots are best examples for AI in robotics, recently the intelligent Humanoid robot named as Erica and Sophia has been developed which can talk and behave like humans.

10. AI in Entertainment:

- We are currently using some AI based applications in our daily life with some entertainment services such as Netflix or Amazon. With the help of ML/AI algorithms, these services show the recommendations for programs or shows.

11. AI in Agriculture:

- Agriculture is an area which requires various resources, labor, money, and time for best result. Now a day's agriculture is becoming digital, and AI is emerging in this field. Agriculture is applying AI as agriculture robotics, solid and crop monitoring, predictive analysis. AI in agriculture can be very helpful for farmers.

12. AI in E-commerce:

- AI is providing a competitive edge to the e-commerce industry, and it is becoming more demanding in the e-commerce business. AI is helping shoppers to discover associated products with recommended size, color, or even brand.

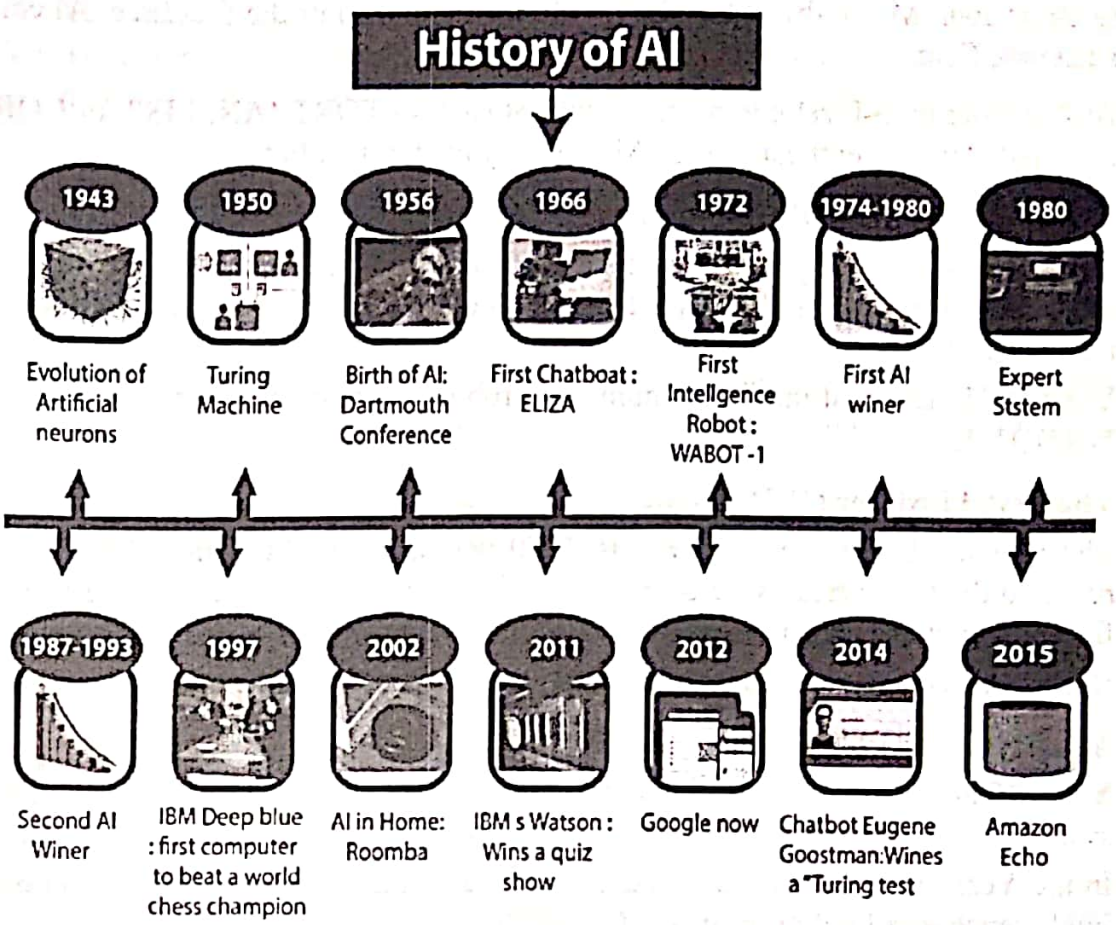
13. AI in education:

- AI can automate grading so that the tutor can have more time to teach. AI chatbot can communicate with students as a teaching assistant.
- AI in the future can be work as a personal virtual tutor for students, which will be accessible easily at any time and any place.

➢ History of Artificial Intelligence:

Artificial Intelligence is not a new word and not a new technology for researchers. This technology is much older than you would imagine. Even there are the myths of Mechanical men in Ancient Greek and Egyptian Myths. Following are some milestones in the history of AI which defines the journey from the AI generation to till date development.

Scanned with CamScanner



- **Maturation of Artificial Intelligence (1943-1952):**
  - **Year 1943:** The first work which is now recognized as AI was done by Warren McCulloch and Walter pits in 1943. They proposed a model of **artificial neurons.**
  - **Year 1949:** Donald Hebb demonstrated an updating rule for modifying the connection strength between neurons. His rule is now called **Hebbian learning.**
  - **Year 1950:** The Alan Turing who was an English mathematician and pioneered Machine learning in 1950. Alan Turing publishes **"Computing Machinery and Intelligence"** in which he proposed a test. The test can check the machine's ability to exhibit intelligent behavior equivalent to human intelligence, called a **Turing test.**
- **The birth of Artificial Intelligence (1952-1956):**
  - **Year 1955:** An Allen Newell and Herbert A. Simon created the "first artificial intelligence program"Which was named as **"Logic Theorist".** This program had proved 38 of 52 Mathematics theorems, and find new and more elegant proofs for some theorems.

Scanned with CamScanner

- **Year 1956:** The word "Artificial Intelligence" first adopted by American Computer scientist John McCarthy at the Dartmouth Conference. For the first time, AI coined as an academic field.

  At that time high-level computer languages such as FORTRAN, LISP, or COBOL were invented. And the enthusiasm for AI was very high at that time.

- **The golden years-Early enthusiasm (1956-1974):**
- **Year 1966:** The researchers emphasized developing algorithms which can solve mathematical problems. Joseph Weizenbaum created the first chatbot in 1966, which was named as ELIZA.
- **Year 1972:** The first intelligent humanoid robot was built in Japan which was named as WABOT-1.

- **The first AI winter (1974-1980):**
- The duration between years 1974 to 1980 was the first AI winter duration. AI winter refers to the time period where computer scientist dealt with a severe shortage of funding from government for AI researches.
- During AI winters, an interest of publicity on artificial intelligence was decreased.

- **A boom of AI (1980-1987):**
- **Year 1980:** After AI winter duration, AI came back with "Expert System". Expert systems were programmed that emulate the decision-making ability of a human expert.
- In the Year 1980, the first national conference of the American Association of Artificial Intelligence was held at Stanford University.

- **The second AI winter (1987-1993):**
- The duration between the years 1987 to 1993 was the second AI Winter duration.
- Again Investors and government stopped in funding for AI research as due to high cost but not efficient result. The expert system such as XCON was very cost effective.

- **The emergence of intelligent agents (1993-2011):**
- **Year 1997:** In the year 1997, IBM Deep Blue beats world chess champion, Gary Kasparov, and became the first computer to beat a world chess champion.
- **Year 2002:** for the first time, AI entered the home in the form of Roomba, a vacuum cleaner.
- **Year 2006:** AI came in the Business world till the year 2006. Companies like Facebook, Twitter, and Netflix also started using AI.

- **Deep learning, big data and artificial general intelligence (2011-present):**

Scanned with CamScanner

- **Year 2011:** In the year 2011, IBM's Watson won jeopardy, a quiz show, where it had to solve the complex questions as well as riddles. Watson had proved that it could understand natural language and can solve tricky questions quickly.
- **Year 2012:** Google has launched an Android app feature "Google now", which was able to provide information to the user as a prediction.
- **Year 2014:** In the year 2014, Chatbot "Eugene Goostman" won a competition in the infamous "Turing test."
- **Year 2018:** The "Project Debater" from IBM debated on complex topics with two master debaters and also performed extremely well.
- Google has demonstrated an AI program "Duplex" which was a virtual assistant and which had taken hairdresser appointment on call, and lady on other side didn't notice that she was talking with the machine.

Now AI has developed to a remarkable level. The concept of Deep learning, big data, and data science are now trending like a boom. Nowadays companies like Google, Facebook, IBM, and Amazon are working with AI and creating amazing devices. The future of Artificial Intelligence is inspiring and will come with high intelligence.

1.3 Agent Environment in AI:

An environment is everything in the world which surrounds the agent, but it is not a part of an agent itself. An environment can be described as a situation in which an agent is present.

The environment is where agent lives, operate and provide the agent with something to sense and act upon it. An environment is mostly said to be non-feministic.

➢ Features of Environment:

As per Russell and Norvig, an environment can have various features from the point of view of an agent:

1. Fully observable vs Partially Observable
2. Static vs Dynamic
3. Discrete vs Continuous
4. Deterministic vs Stochastic
5. Single-agent vs Multi-agent
6. Episodic vs sequential
7. Known vs Unknown
8. Accessible vs Inaccessible

1. Fully observable vs Partially Observable:

- If an agent sensor can sense or access the complete state of an environment at each point of time then it is **a fully observable** environment, else it is **partially observable.**

Scanned with CamScanner

- A fully observable environment is easy as there is no need to maintain the internal state to keep track history of the world.
- An agent with no sensors in all environments then such an environment is called as **unobservable.**

2. Deterministic vs Stochastic:

- If an agent's current state and selected action can completely determine the next state of the environment, then such environment is called a deterministic environment.
- A stochastic environment is random in nature and cannot be determined completely by an agent.
- In a deterministic, fully observable environment, agent does not need to worry about uncertainty.

3. Episodic vs Sequential:

- In an episodic environment, there is a series of one-shot actions, and only the current percept is required for the action.
- However, in Sequential environment, an agent requires memory of past actions to determine the next best actions.

4. Single-agent vs Multi-agent:

- If only one agent is involved in an environment, and operating by itself then such an environment is called single agent environment.
- However, if multiple agents are operating in an environment, then such an environment is called a multi-agent environment.
- The agent design problems in the multi-agent environment are different from single agent environment.

5. Static vs Dynamic:

- If the environment can change itself while an agent is deliberating then such environment is called a dynamic environment else it is called a static environment.
- Static environments are easy to deal because an agent does not need to continue looking at the world while deciding for an action.
- However for dynamic environment, agents need to keep looking at the world at each action.
- Taxi driving is an example of a dynamic environment whereas Crossword puzzles are an example of a static environment.

6. Discrete vs Continuous:

- If in an environment there are a finite number of percepts and actions that can be performed within it, then such an environment is called a discrete environment else it is called continuous environment.

Scanned with CamScanner

- A chess gamecomes under discrete environment as there is a finite number of moves that can be performed.
- A self-driving car is an example of a continuous environment.

7. Known vs Unknown:

- Known and unknown are not actually a feature of an environment, but it is an agent's state of knowledge to perform an action.
- In a known environment, the results for all actions are known to the agent. While in unknown environment, agent needs to learn how it works in order to perform an action.
- It is quite possible that a known environment to be partially observable and an Unknown environment to be fully observable.

8. Accessible vs Inaccessible:

- If an agent can obtain complete and accurate information about the state's environment, then such an environment is called an Accessible environment else it is called inaccessible.
- An empty room whose state can be defined by its temperature is an example of an accessible environment.

Information about an event on earth is an example of Inaccessible environment.

1.4 The Concept of Rationality:

❖ Rationality:

Rationality is nothing but status of being reasonable, sensible, and having good sense of judgment.

Rationality is concerned with expected actions and results depending upon what the agent has perceived. Performing actions with the aim of obtaining useful information is an important part of rationality.

❖ What is Ideal Rational Agent?

An ideal rational agent is the one, which is capable of doing expected actions to maximize its performance measure, on the basis of –

- Its percept sequence
- Its built-in knowledge base

Rationality of an agent depends on the following –

- The **performance measures**, which determine the degree of success.
- Agent's **Percept Sequence** till now.
- The agent's **prior knowledge about the environment.**
- The **actions** that the agent can carry out.

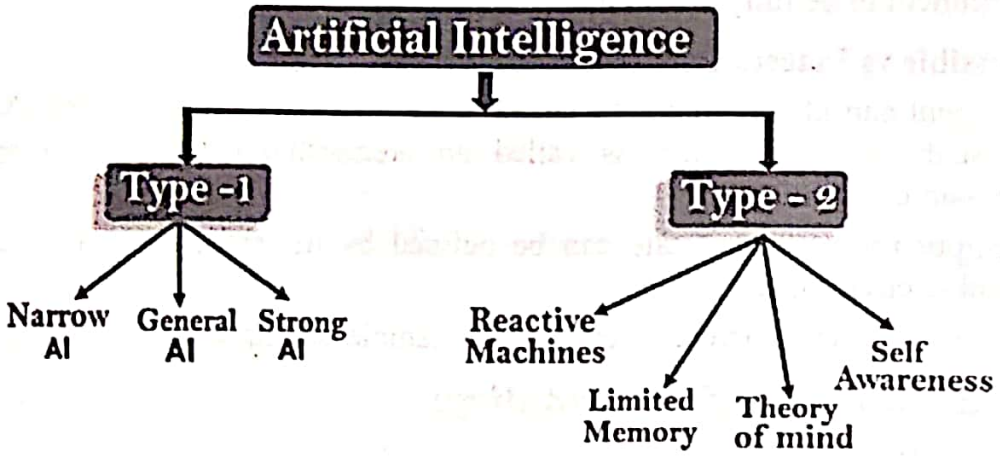
Scanned with CamScanner

A rational agent always performs right action, where the right action means the action that causes the agent to be most successful in the given percept sequence. The problem the agent solves is characterized by Performance Measure, Environment, Actuators, and Sensors (PEAS).

1.5 Nature of Environment:

❖ Types of Artificial Intelligence:

Artificial Intelligence can be divided in various types, there are mainly two types of main categorization which are based on capabilities and based on functionally of AI. Following is flow diagram which explain the types of AI.



➢ AI type-1: Based on Capabilities:

1. Weak AI or Narrow AI:

- Narrow AI is a type of AI which is able to perform a dedicated task with intelligence.The most common and currently available AI is Narrow AI in the world of Artificial Intelligence.
- Narrow AI cannot perform beyond its field or limitations, as it is only trained for one specific task. Hence it is also termed as weak AI. Narrow AI can fail in unpredictable ways if it goes beyond its limits.
- Apple Siriis a good example of Narrow AI, but it operates with a limited pre-defined range of functions.
- IBM's Watson supercomputer also comes under Narrow AI, as it uses an Expert system approach combined with Machine learning and natural language processing.
- Some Examples of Narrow AI are playing chess, purchasing suggestions on e-commerce site, self-driving cars, speech recognition, and image recognition.

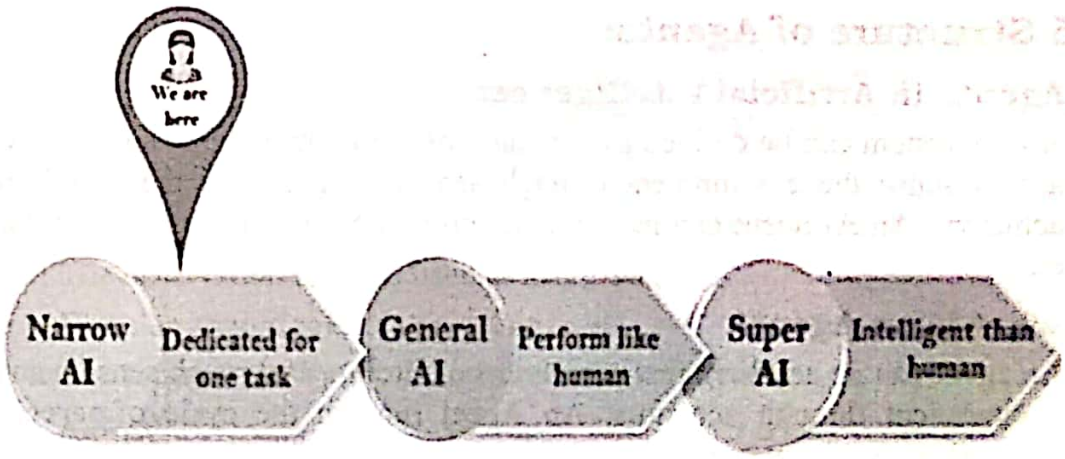
Scanned with CamScanner

2. General AI:

- General AI is a type of intelligence which could perform any intellectual task with efficiency like a human.
- The idea behind the general AI to make such a system which could be smarter and think like a human by its own.
- Currently, there is no such system exist which could come under general AI and can perform any task as perfect as a human.
- The worldwide researchers are now focused on developing machines with General AI.
- As systems with general AI are still under research, and it will take lots of efforts and time to develop such systems.

3. Super AI:

- Super AI is a level of Intelligence of Systems at which machines could surpass human intelligence, and can perform any task better than human with cognitive properties. It is an outcome of general AI.
- Some key characteristics of strong AI include capability include the ability to think, to reason,solve the puzzle, make judgments, plan, learn, and communicate by its own.
- Super AI is still a hypothetical concept of Artificial Intelligence. Development of such systems in real is still world changing task.



➢ **Artificial Intelligence type-2: Based on functionality**

1. Reactive Machines:

- Purely reactive machines are the most basic types of Artificial Intelligence.
- Such AI systems do not store memories or past experiences for future actions.
- These machines only focus on current scenarios and react on it as per possible best action.
- IBM's Deep Blue system is an example of reactive machines.
- Google's AlphaGo is also an example of reactive machines.

Scanned with CamScanner

2. **Limited Memory:**

- Limited memory machines can store past experiences or some data for a short period of time.
- These machines can use stored data for a limited time period only.
- Self-driving cars are one of the best examples of Limited Memory systems. These cars can store recent speed of nearby cars, the distance of other cars, speed limit, and other information to navigate the road.

3. **Theory of Mind:**

- Theory of Mind AI should understand the human emotions, people, beliefs, and be able to interact socially like humans.
- This type of AI machines are still not developed, but researchers are making lots of efforts and improvement for developing such AI machines.

4. **Self-Awareness:**

- Self-awareness AI is the future of Artificial Intelligence. These machines will be super intelligent, and will have their own consciousness, sentiments, and self-awareness.
- These machines will be smarter than human mind.
- Self-Awareness AI does not exist in reality still and it is a hypothetical concept.

1.6 Structure of Agents:

❖ Agents in Artificial Intelligence:

An AI system can be defined as the study of the rational agent and its environment. The agents sense the environment through sensors and act on their environment through actuators. An AI agent can have mental properties such as knowledge, belief, intention, etc.

➢ What is an Agent?

An agent can be anything that perceiveits environment through sensors and act upon that environment through actuators. An Agent runs in the cycle of **perceiving**, **thinking**, and **acting**. An agent can be:

- **Human-Agent:** A human agent has eyes, ears, and other organs which work for sensors and hand, legs, vocal tract work for actuators.
- **Robotic Agent:** A robotic agent can have cameras, infrared range finder, NLP for sensors and various motors for actuators.
- **Software Agent:** Software agent can have keystrokes, file contents as sensory input and act on those inputs and display output on the screen.

Hence the world around us is full of agents such as thermostat, cellphone, camera, and even we are also agents.

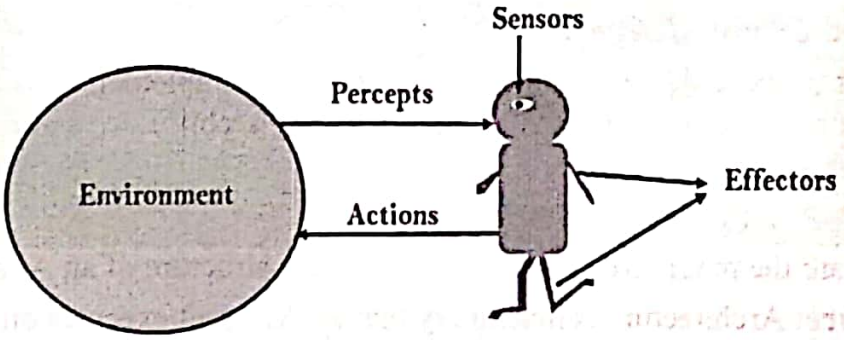
Scanned with CamScanner

Before moving forward, we should first know about sensors, effectors, and actuators.

Sensor: Sensor is a device which detects the change in the environment and sends the information to other electronic devices. An agent observes its environment through sensors.

Actuators: Actuators are the component of machines that converts energy into motion. The actuators are only responsible for moving and controlling a system. An actuator can be an electric motor, gears, rails, etc.

Effectors: Effectors are the devices which affect the environment. Effectors can be legs, wheels, arms, fingers, wings, fins, and display screen.



➢ Intelligent Agents:

An intelligent agent is an autonomous entity which act upon an environment using sensors and actuators for achieving goals. An intelligent agent may learn from the environment to achieve their goals. A thermostat is an example of an intelligent agent.

Following are the main four rules for an AI agent:

- **Rule 1:** An AI agent must have the ability to perceive the environment.
- **Rule 2:** The observation must be used to make decisions.
- **Rule 3:** Decision should result in an action.
- **Rule 4:** The action taken by an AI agent must be a rational action.

➢ Rational Agent:

A rational agent is an agent which has clear preference, models uncertainty, and acts in a way to maximize its performance measure with all possible actions.

A rational agent is said to perform the right things. AI is about creating rational agents to use for game theory and decision theory for various real-world scenarios.

For an AI agent, the rational action is most important because in AI reinforcement learning algorithm, for each best possible action, agent gets the positive reward and for each wrong action, an agent gets a negative reward.

Note: Rational agents in AI are very similar to intelligent agents.

Scanned with CamScanner

➢ **Rationality:**

The rationality of an agent is measured by its performance measure. Rationality can be judged on the basis of following points:

- Performance measure which defines the success criterion.
- Agent prior knowledge of its environment.
- Best possible actions that an agent can perform.
- The sequence of percepts.

Note: Rationality differs from Omniscience because an Omniscient agent knows the actual outcome of its action and act accordingly, which is not possible in reality.

❖ **Structure of an AI Agent:**

The task of AI is to design an agent program which implements the agent function. The structure of an intelligent agent is a combination of architecture and agent program. It can be viewed as:

1. Agent = Architecture + Agent program

Following are the main three terms involved in the structure of an AI agent:

Architecture: Architecture is machinery that an AI agent executes on.

Agent Function: Agent function is used to map a percept to an action.

1. $f:P^* \rightarrow A$

Agent program: Agent program is an implementation of agent function. An agent program executes on the physical architecture to produce function f.

The environment is the **Task Environment (problem)** for which the **Rational Agent is the solution.** Any task environment is characterised on the basis of PEAS.

PEAS is a type of model on which an AI agent works upon. When we define an AI agent or rational agent, then we can group its properties under PEAS representation model. It is made up of four words:

- **P:** Performance measure
- **E:** Environment
- **A:** Actuators
- **S:** Sensors

Here performance measure is the objective for the success of an agent's behavior.

1. **Performance** – What is the performance characteristic which would either make the agent successful or not. For example, as per the previous example clean floor, optimal energy consumption might be performance measures.
2. **Environment** – Physical characteristics and constraints expected. For example, wood floors, furniture in the way etc.

Scanned with CamScanner

3. **Actuators** – The physical or logical constructs which would take action. For example for the vacuum cleaner, these are the suction pumps.
4. **Sensors** – Again physical or logical constructs which would sense the environment. From our previous example, these are cameras and dirt sensors.

❖ **PEAS Representation:**

PEAS for self-driving cars:

Let's suppose a self-driving car then PEAS representation will be:

Performance: Safety, time, legal drive, comfort

Environment: Roads, other vehicles, road signs, pedestrian

Actuators: Steering, accelerator, brake, signal, horn

Sensors: Camera, GPS, speedometer, odometer, accelerometer, sonar.

❖ **Example of Agents with their PEAS representation:**

Agent	Performance measure	Environment	Actuators	Sensors
1. Medical Diagnose	○ Healthy patient ○ Minimized cost	○ Patient ○ Hospital ○ Staff	○ Tests ○ Treatments	○ Keyboard (Entry of symptoms)
2. Vacuum Cleaner	○ Cleanness ○ Efficiency ○ Battery life ○ Security	○ Room ○ Table ○ Wood floor ○ Carpet ○ Various obstacles	○ Wheels ○ Brushes ○ Vacuum Extractor	○ Camera ○ Dirt detection sensor ○ Cliff sensor ○ Bump Sensor ○ Infrared Wall Sensor
3. Part - Picking Robot	○ Percentage of parts in correct bins.	○ Conveyor belt with parts, ○ Bins	○ Jointed Arms ○ Hand	○ Camera ○ Joint angle sensors.

Scanned with CamScanner

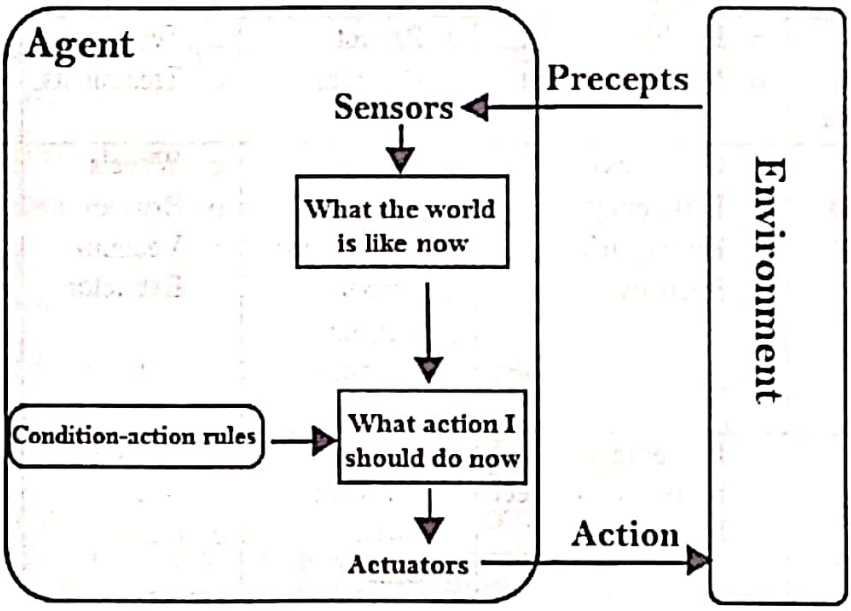
❖ Types of AI Agents:

Agents can be grouped into five classes based on their degree of perceived intelligence and capability. All these agents can improve their performance and generate better action over the time. These are given below:

1) Simple Reflex Agent
2) Model-based reflex agent
3) Goal-based agents
4) Utility-based agent
5) Learning agent

1. Simple Reflex agent:

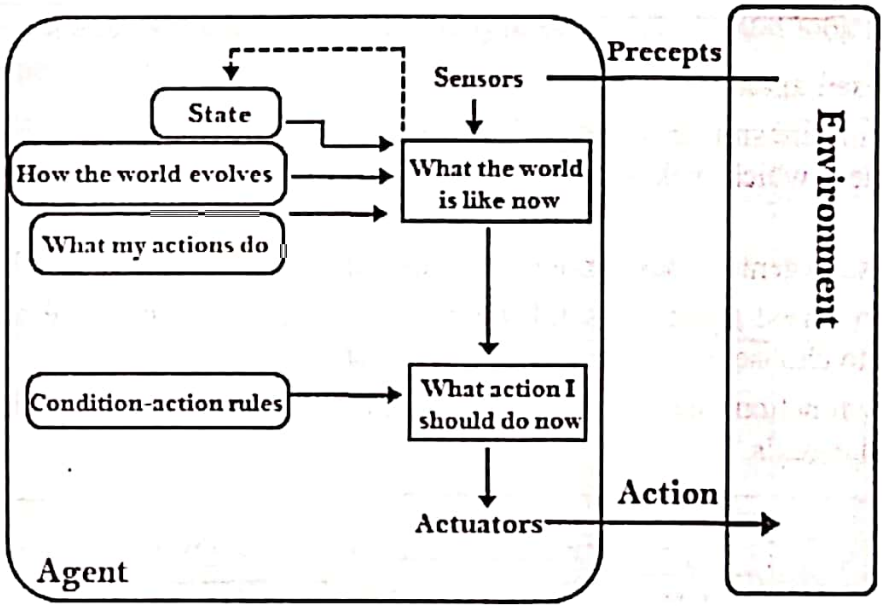
- The Simple reflex agents are the simplest agents. These agents take decisions on the basis of the current percepts and ignore the rest of the percept history.
- These agents only succeed in the fully observable environment.
- The Simple reflex agent does not consider any part of percepts history during their decision and action process.
- The Simple reflex agent works on Condition-action rule, which means it maps the current state to action. Such as a Room Cleaner agent, it works only if there is dirt in the room.
- Problems for the simple reflex agent design approach:
- They have very limited intelligence
- They do not have knowledge of non-perceptual parts of the current state
- Mostly too big to generate and to store.
- Not adaptive to changes in the environment.



Scanned with CamScanner

2. Model-based reflex agent:

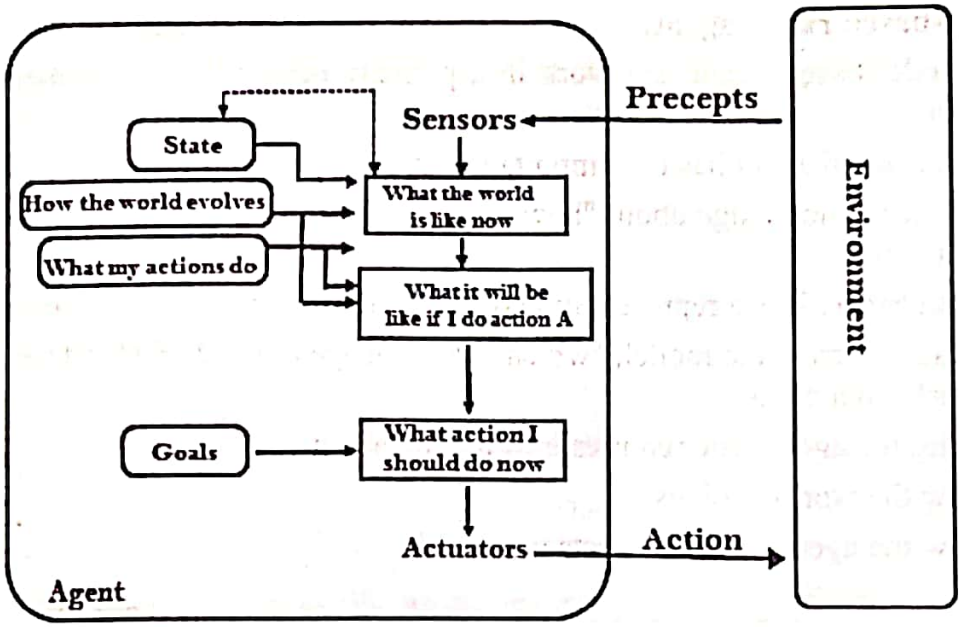
- The Model-based agent can work in a partially observable environment, and track the situation.
- A model-based agent has two important factors:
- **Model:** It is knowledge about "how things happen in the world," so it is called a Model-based agent.
- **Internal State:** It is a representation of the current state based on percept history.
- These agents have the model, "which is knowledge of the world" and based on the model they perform actions.
- Updating the agent state requires information about:
  a. How the world evolves
  b. How the agent's action affects the world.



3. Goal-based agents:

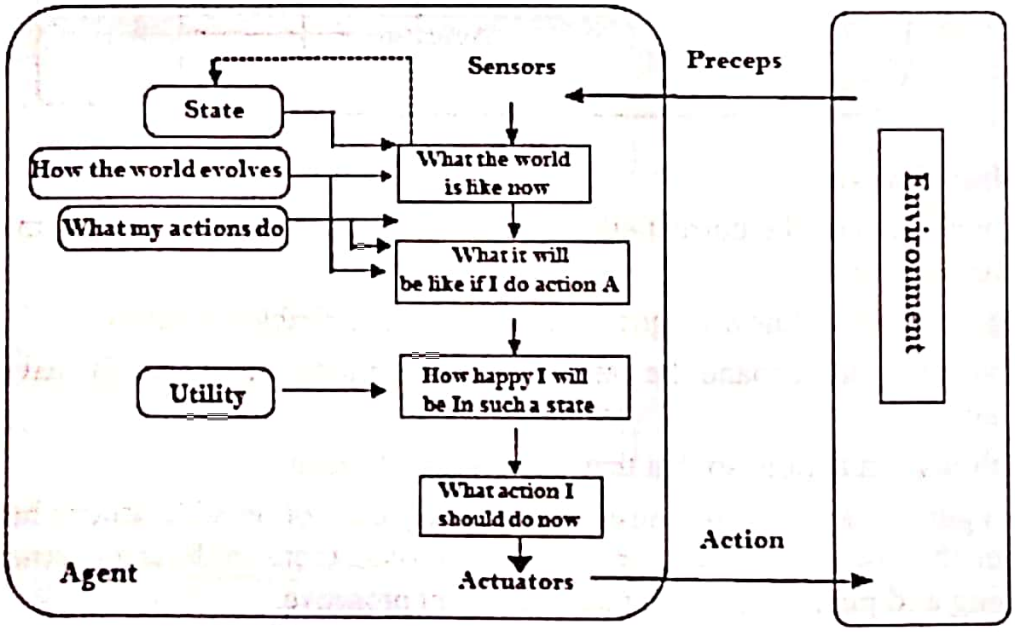
- The knowledge of the current state environment is not always sufficient to decide for an agent to what to do.
- The agent needs to know its goal which describes desirable situations.
- Goal-based agents expand the capabilities of the model-based agent by having the "goal" information.
- They choose an action, so that they can achieve the goal.
- These agents may have to consider a long sequence of possible actions before deciding whether the goal is achieved or not. Such considerations of different scenario are called searching and planning, which makes an agent proactive.

Scanned with CamScanner



4. Utility-based agents:

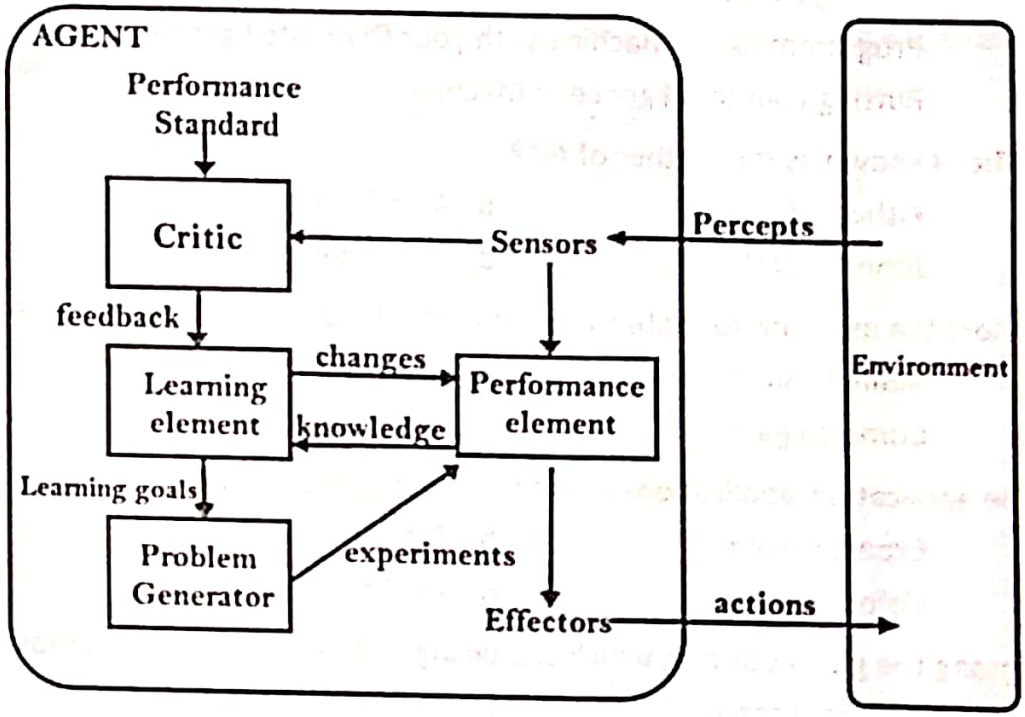
- These agents are similar to the goal-based agent but provide an extra component of utility measurement which makes them different by providing a measure of success at a given state.
- Utility-based agent act based not only goals but also the best way to achieve the goal.
- The Utility-based agent is useful when there are multiple possible alternatives, and an agent has to choose in order to perform the best action.
- The utility function maps each state to a real number to check how efficiently each action achieves the goals.



Scanned with CamScanner

5. Learning Agents:

- A learning agent in AI is the type of agent which can learn from its past experiences, or it has learning capabilities.
- It starts to act with basic knowledge and then able to act and adapt automatically through learning.
- A learning agent has mainly four conceptual components, which are:
  a. **Learning element:** It is responsible for making improvements by learning from environment
  b. **Critic:** Learning element takes feedback from critic which describes that how well the agent is doing with respect to a fixed performance standard.
  c. **Performance element:** It is responsible for selecting external action
  d. **Problem generator:** This component is responsible for suggesting actions that will lead to new and informative experiences.
- Hence, learning agents are able to learn, analyze performance, and look for new ways to improve the performance.



Scanned with CamScanner

Exercises

❖ Answer the following Questions in brief.

1. Write all the eight, different definitions of Artificial Intelligence.
2. Short note on State of the Art applications of Artificial Intelligence.
3. Short note on task environment. Also explain PEAS description.
4. Write a short note on Agents.

❖ Multipal choice Question - MCQs:

1. Artificial Intelligence is about _____.
   a. Playing a game on Computer
   b. Making a machine Intelligent
   c. Programming on Machine with your Own Intelligence
   d. Putting your intelligence in Machine
2. Who is known as the -Father of Al"?
   a. Fisher Ada
   b. Alan Turing
   c. John McCarthy
   d. Allen Newell
3. 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
4. The application/applications of Artificial Intelligence is/are
   a. Expert Systems
   b. Gaming
   c. Vision Systems
   d. All of the above
5. Among the given options, which search algorithm requires less memory?
   a. Optimal Search
   b. Depth First Search
   c. Breadth-First Search
   d. Linear Search
6. 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

Scanned with CamScanner

7. A technique that was developed to determine whether a machine could or could not demonstrate the artificial intelligence known as the _____.
   a. Boolean Algebra b. Turing Test
   c. Logarithm d. Algorithm

8. The component of an Expert system is _____.
   a. Knowledge Base b. Inference Engine
   c. User Interface d. All of the above

9. Which algorithm is used in the Game tree to make decisions of Win/Lose?
   a. Heuristic Search Algorithm b. DFS/BFS algorithm
   c. Greedy Search Algorithm d. Min/Max algorithm

10. An AI agent perceives and acts upon the environment using _____.
   a. Sensors b. Perceiver
   c. Actuators d. Both a and c

Answers:

1. a	2. b	3. b	4. d	5. b
6. d	7. b	8. d	9. d	10. d

Scanned with CamScanner

UNIT-2

Problem Solving by Searching

- Problem Solving by searching
- Problem-Solving Agents
  - Well defined problem and solutions
  - Formulating problems
- Example Problems
  - Toy problems
- Searching for Solution
- Uninformed Search Strategies
  - Concept of BFS
  - Concept of DFS
  - Depth-limited search
  - Iterative deepening DFS
  - Bidirectional search
- Informed (Heuristic) Search Strategies
  - Concept of Greedy BFS
  - A* search: Minimizing the total estimated solution cost
- Case Study: Applications of AI in transportation.

Scanned with CamScanner

Unit –2 Problem Solving by Searching

2.1 Problem Solving by Searching:

In Artificial Intelligence searching method is the best technique to solve the problems whether it is defined or not defined. AI used search strategies to solve the specific problem.

To solve the problems in Artificial Intelligence it need following things:

a. Problem Define
b. Problem Analyze
c. Isolate problem and knowledge of task
d. Get techniques
e. Choose techniques to solve problems

To build a system to solve a particular problem first we need to specify the initial system (I) what exactly the problem is and the acceptable final solutions to the problem. Isolate the problem and read about the problems then get multiple possible techniques that are necessary to solve the problem, after that we can get the best technique and apply it (them) to the particular problem.

Define the problem as a **state space search.** All the possible state for solving any particular problem is called as **state space.** For searching the particular state from the state space we use state space search. The structure of state space is of two types:

- In some problems set of states and rules are already defined and well structured organize. We have to convert some given condition into the required condition using a set of operations which are already decided.
- In some problems these kinds of representation occurs naturally and are not well structured. In this we have to make the desired structure with the help of some techniques to find the path from the initial state to goal state. Search is an important procedure for the problems to find the, to and Forth solution no direct techniques are used in it.

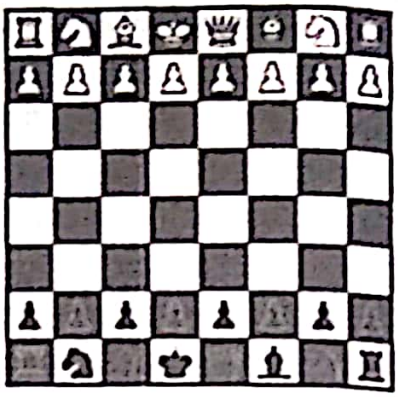
2.2 Problem solving Agents:

Problem solving agents are the goal based agents which are also called as **problem solving agents.** Simple Problem solving agents have limited strategies to solve the problems. They solve the problems on current perceptions they don't know what they want and what are the goals. Problem solving agents are such that they create the sequential environment to maximize the performance of searching method to solve any problem. Let illustrate the problems **play chess game** and **milk conman problem.**

Scanned with CamScanner

❖ Play chess Problem:

We have to build a program which can play chess. Here the required conditions are already defined which can be the position of the players, moves, operations, initial state and goal state, legal game to win and lost rules.



In the **play chess game** the rules are already provided and the starting position is defines as an 8 by 8 position where each box contains its value and has some opening rules with the legal move. There are many ways to apply this rules because of that it is too difficult and lengthy one could not get the goal without any mistakes. To solve this problem if possible one should look for a way to write the rules with legal moves.

❖ Milk Conman Problem:

We are given two conman (utensil to pour liquid), a 6 liter and 4 liter. We don't have measuring cup or any mark. We are given a tap to fill the Conman. How can we get exact 2liter of milk into the 6 liter of Conman?

One solution of milk Conman problem	
Liters in the 6 liter of Conman	Liters in the 4 liter of Conman
0	0
0	4
4	0
4	4
6	2
0	2
2	0

The above two problems can be solve with the strategies we discussed already. Both the problems are different and can be solved by using some rules, strategies, approximation, some combinations, and movements until the goal state is found. Play chess game can be solved by using single state search while for milk conman problem different combinations of searching is being used.

Scanned with CamScanner

Problem Solving Agents	
input:	p, Assumption
static:	s, action sequence, empty state, description of present state,
	g, goal, null
	problem, problem changed
state: ←	UPDATE-STATE (state,p)
if	s is empty then
g -	GOAL, (state) CHANGED
probl em	<- PROBLEM CHANGED (state, g)
s -	SEARCH (problem)
action -	recommendation (s, state)
s ←	REMINDER (s, state)
return	action

❖ Formulating problems:

Formulating problem is depending on the knowledge, environment and perception of the problem. We can divide the problem as-

a. Single state problem
b. Multiple state problem
c. Contingency problem
d. Exploration problem

Let us learn the problem with assumption before that we should understand the problem first, when the initial or present state is known with that one should get to know about the net action, sequence of action or sometimes we get goal state which is called **single state problem**. If all the action are known but because of some rules, we get the limited access of the action then the possibility of happening the particular state is greater than 1, thus it can be known as **multiple state problem.** When there is multiple state problem then to solving the problem the process can also be more than one. Probability can be used to solve single problem thus this type of problem requires the sequence of action taken. It depends on the possible contingency which can also call as **contingency problem**. The real time problems are not exact predicted but some additional information like problem state can be known.

Scanned with CamScanner

2.2.1 Well defined problems and solutions:

Problem is a special part to solve the problem. It is the collection of information that one can decide what exactly we want and what to do next. There are elements of problem: STATE, ACTION, PATH and GOAL.

State can be defined as initial state or end state both. In initial state the problem could be known to be itself. When the problem is already defines b some rules then one could get end state. When the problem is already defined by some rules then one could get end state. After the initial state the set of possible actions available for the problem can be determined. In action state space is used and the path of state space is any sequence of actions which can apply on one state to another. The path is the sequence of any action from the initial state to reach the goal or End state.

Input
datatype: PROBLEM
elements: INITIAL STATE, ACTION, PATH, GOAL, ENDSTATE

2.3 Example problems:

❖ Toy problems:

- 8- puzzle problem-

Start State

1	2	3
4	8	-
7	6	5

Goal State

1	2	3
4	5	6
7	8	-

Down
Up
Left
Right

The 8– Puzzle problem is 3×3 board with 8 tiles and a blank space in which tiles slides from one place to another only where the blank space is available. In the puzzle we have some rules and conditions of moving the tiles in which every tiles should be move only Up, Down, Left, Right. Tiles cannot move diagonally. We can lead following formulation:

- States: A state describes the location of each eight tiles and also of the blank tiles.
- Operators: Up, Down, Left, Right (blank tiles only).
- Start: Describes the initial state.
- Goal: Describes the goal state.

Scanned with CamScanner

- Path cost: Number of steps taken to reach the goal.

Following are the conditions and states to solve the 8-puzzle problem.

Down-

{(1,2,3), (4,8,0), (7,6,5)}

{(1,2,3), (4,8,5), (7,6,0)}

Left-

{(1,2,3), (4,8,5), (7,0,6)}

Up-

{(1,2,3), (4,0,5), (7,8,6)}

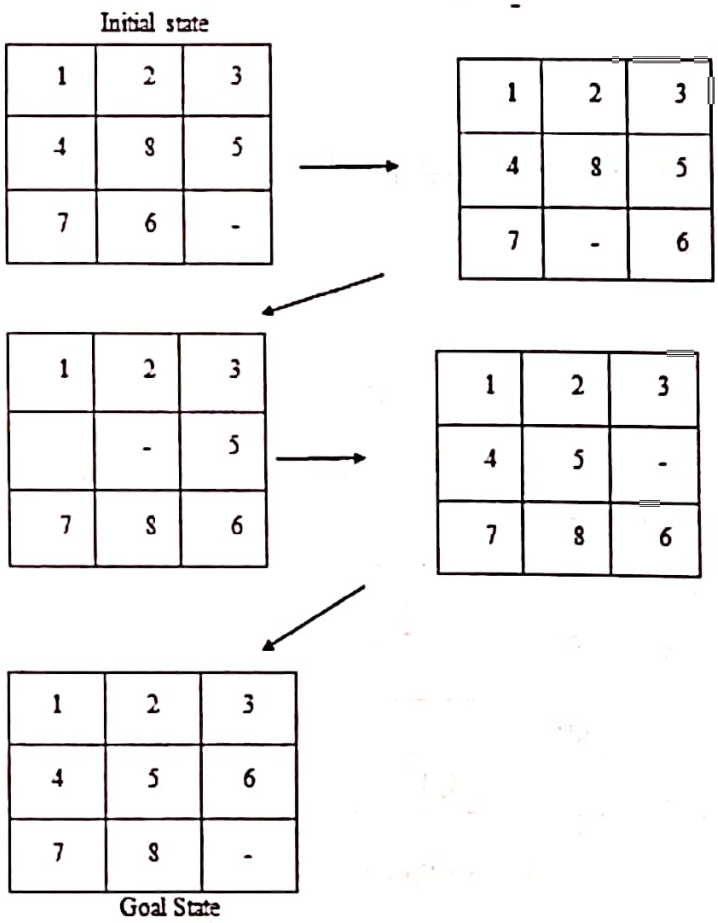
Right-

{(1,2,3), (4,5,0), (7,8,6)}

Down-

{(1,2,3), (4,5,6), (7,8,0)}

Path cost= ("5") i.e. no. of states taken to solve the problem.



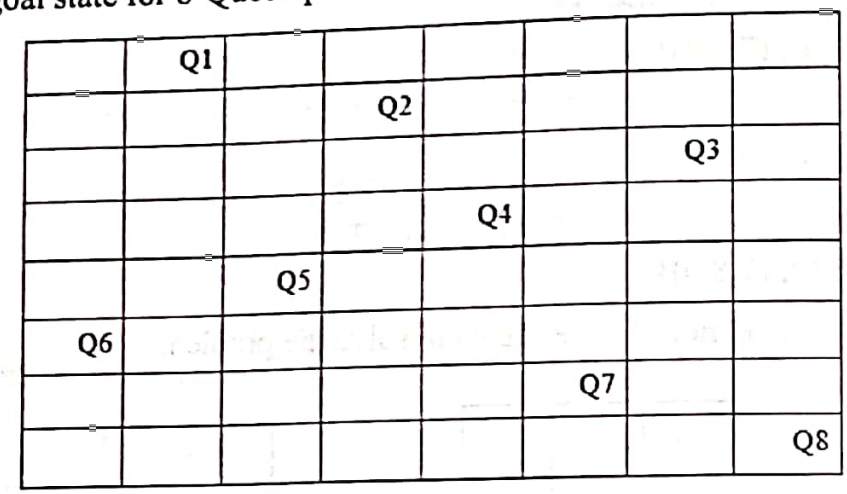
Scanned with CamScanner

❖ **N- Queen Problem:**

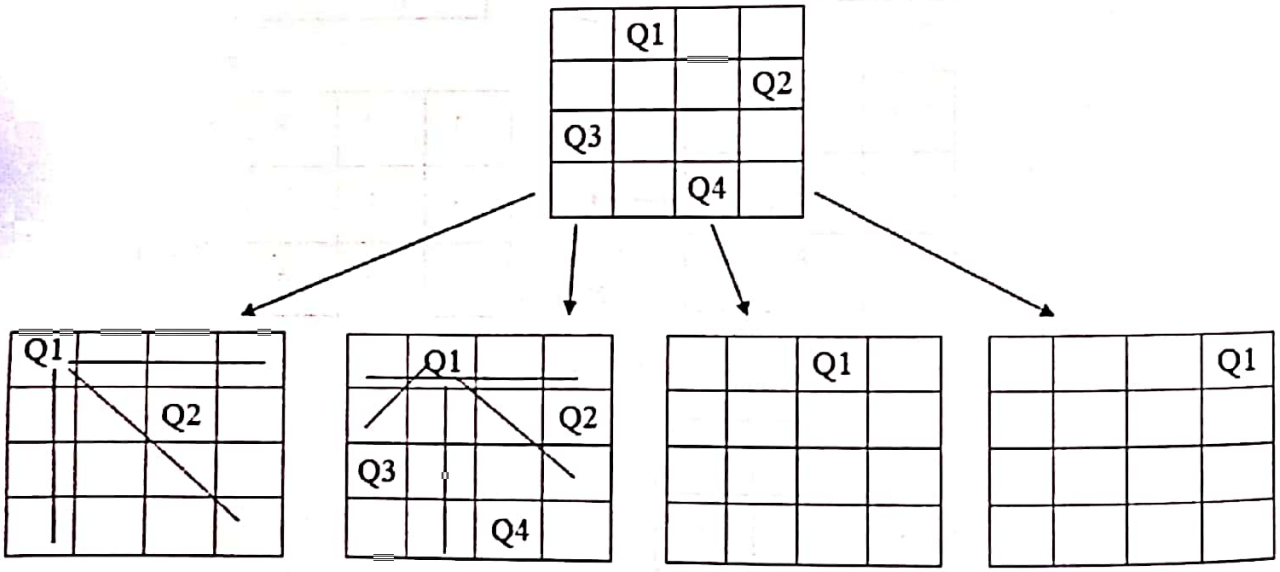
This is the type of constraint satisfactory problem in AI. It has N×N square grid board in which N number of queens is placed. Some constraints/rules are-

1. No row should contain more than one queen.
2. No column should contain more than one queen.
3. No diagonal should contain more than one queen.
4. There should be no row or column without any queen.

Here is the goal state for 8-Queen problem –



Let we discuss and Solving 4-queen problem



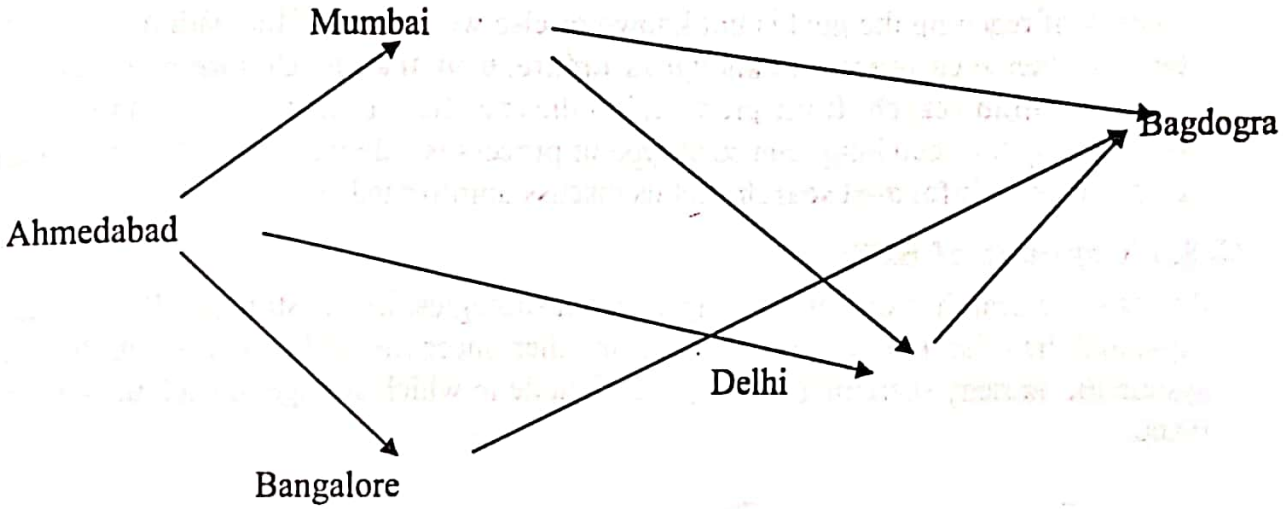
Stop Done (Solution)

Scanned with CamScanner

- State: initial position
- Goal: no queen attack on each other
- Operations: place the queen in an square
- Path cost: zero

❖ **Real world Problems:**

Route finding is defined for specified location, source, destination, path and the links between the states.



Let illustrate from the above diagram a traveler need to go from Ahmedabad to Bagdogra. He has three paths to reach to destination.

Ahmedabad →Delhi → Bagdogra

Ahmedabad →Banglore → Bagdogra

Ahmedabad →Mumbai → Bagdogra

It depends on the path and the distance to reach the destination on less time and low cost. If we talk about road transport then it takes more distance, traffic, low cost of petrol and driving time. We use route finding route finding algorithm on both the road and air transportation. On roads it may be simple but for airline travel route finding algorithm is somehow complex in terms of path, cost, seat availability, quality, time, type of air transport, frequency, mileage and so on. Somehow the actions of this problem are also get known outcomes, flight may be delayed, emergency, maintenance, late or overbooked, fog delayed, connections can also be missed.

2.4 Searching for solution:

We have seen the problem definition, problem types, recognize the solution to finding the solution we used state space and state space search method to find the solution with the path and the sequence to get the goal state. As we discussed above in 4-queens problem

Scanned with CamScanner

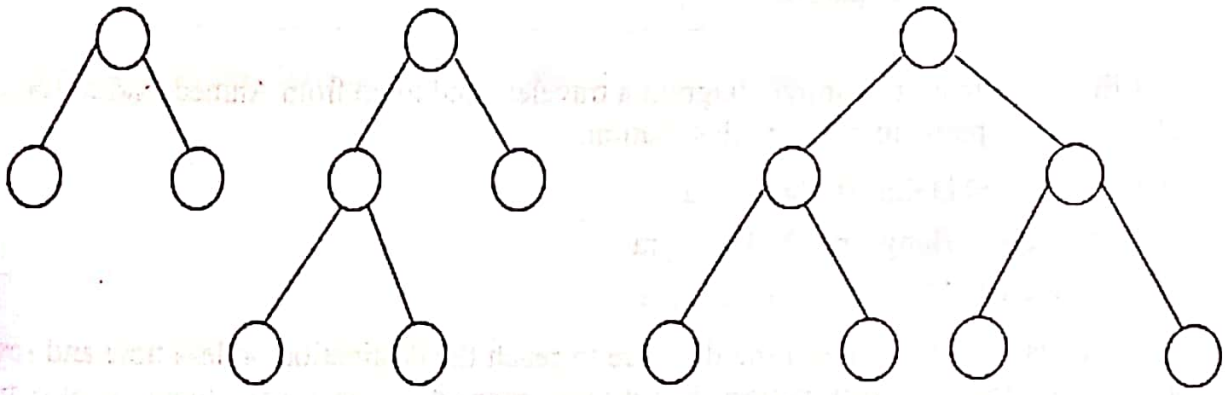
the first method didn't get us the solution so we stopped and try another to get the solution. If the first choice does not lead to the solution we can put that option for later and choose another option to get the solution. We can continue choosing and exploring until the solution is found or the next state is found this process is also known as search strategy. Many search trees are used over state space to find the route and solution and that we will discuss later.

2.5 Uniformed search strategies:

We already discussed about the problems and searching. When the actions and the process of reaching the goal is not known or else we can say if the path of the goal is not defined then such process is known as **uninformed search**. Uninformed search is also called as **blind search**. If the problem has the specific information about the next step or path to help the searching then such type of process is called as **heuristic search** and it is also named as **informed search**. Let us discuss uninformed search.

2.5.1 Concept of BFS:

Breadth first search is one of the simple search strategies. In this strategy all the nodes are expanded, first the root node is expanded and then successor and similar so on. It is very systematic strategy searching occurs on each node in which storage of each node takes place.



Circle represents nodes (0, 1, 2, 3, 4, 5, 6, 7, 8.... so on)

- **Levels of BFS:**

Function BFS (problem)
If Goal
Then Return, Failure
Or else return search (problem, Goal)
Return

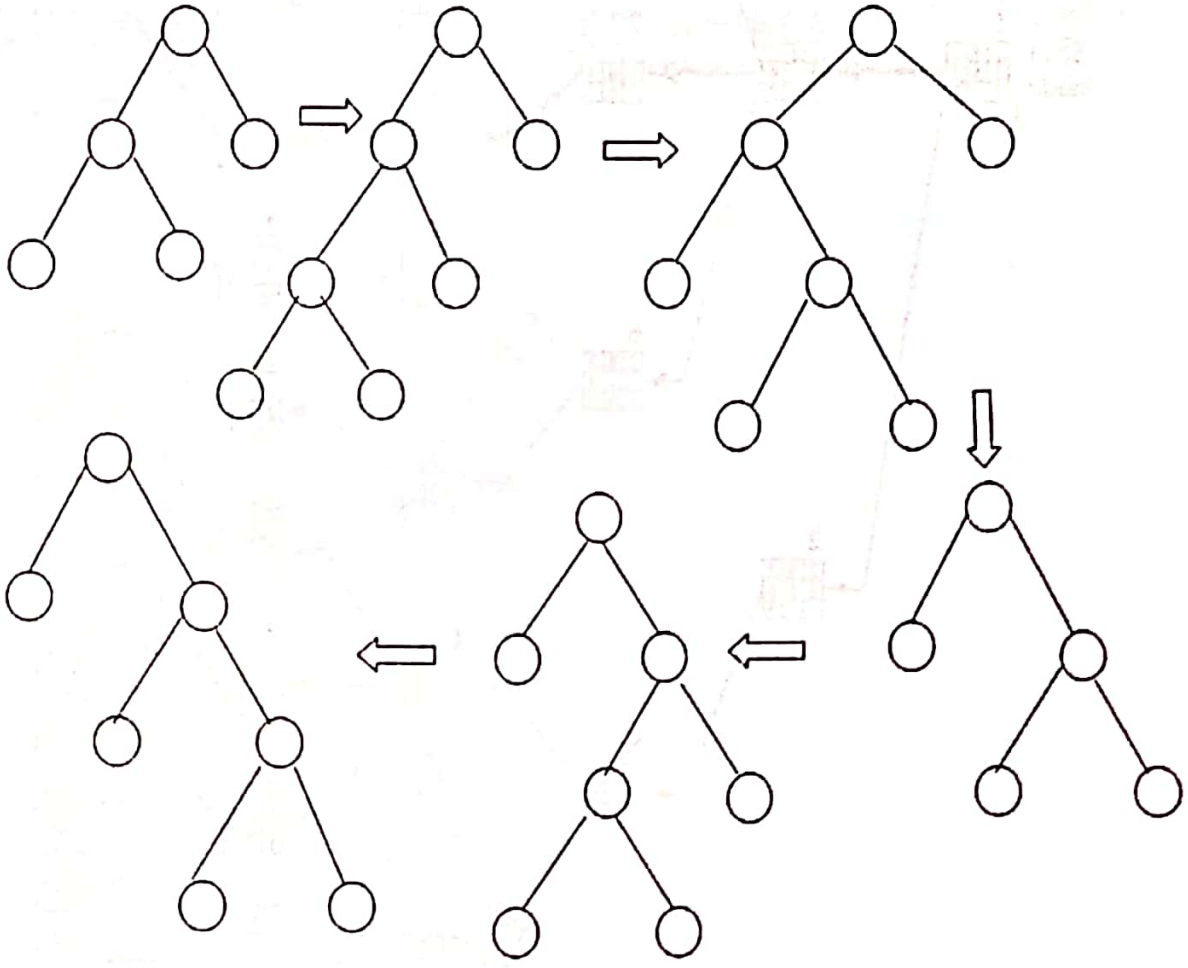
Scanned with CamScanner

If there is a solution for a problem then BFS always find not so deep goal state first, then the second one and so on. But it takes so much of memory and time sometime BFS is not been chosen strategy among all. Because of the searching of every single node, BFS takes time and also memory. Assume 100 nodes can be checked in a second and one node requires 10 bytes of storage thus the more memory and time required which is problem for BFS. To solve any random problem using BFS one could wait for less than 10 minutes and because of the more memory the cost also increase to solve the problem and to reach out to the Goal state.

2.5.2 Concept of DFS:

Depth first search is also known as backtracking search. DFS expands on the nodes, ones it expand on one node it goes on deepest level of the tree only when the search get on the endpoint then it go back to the initial position and so forth.

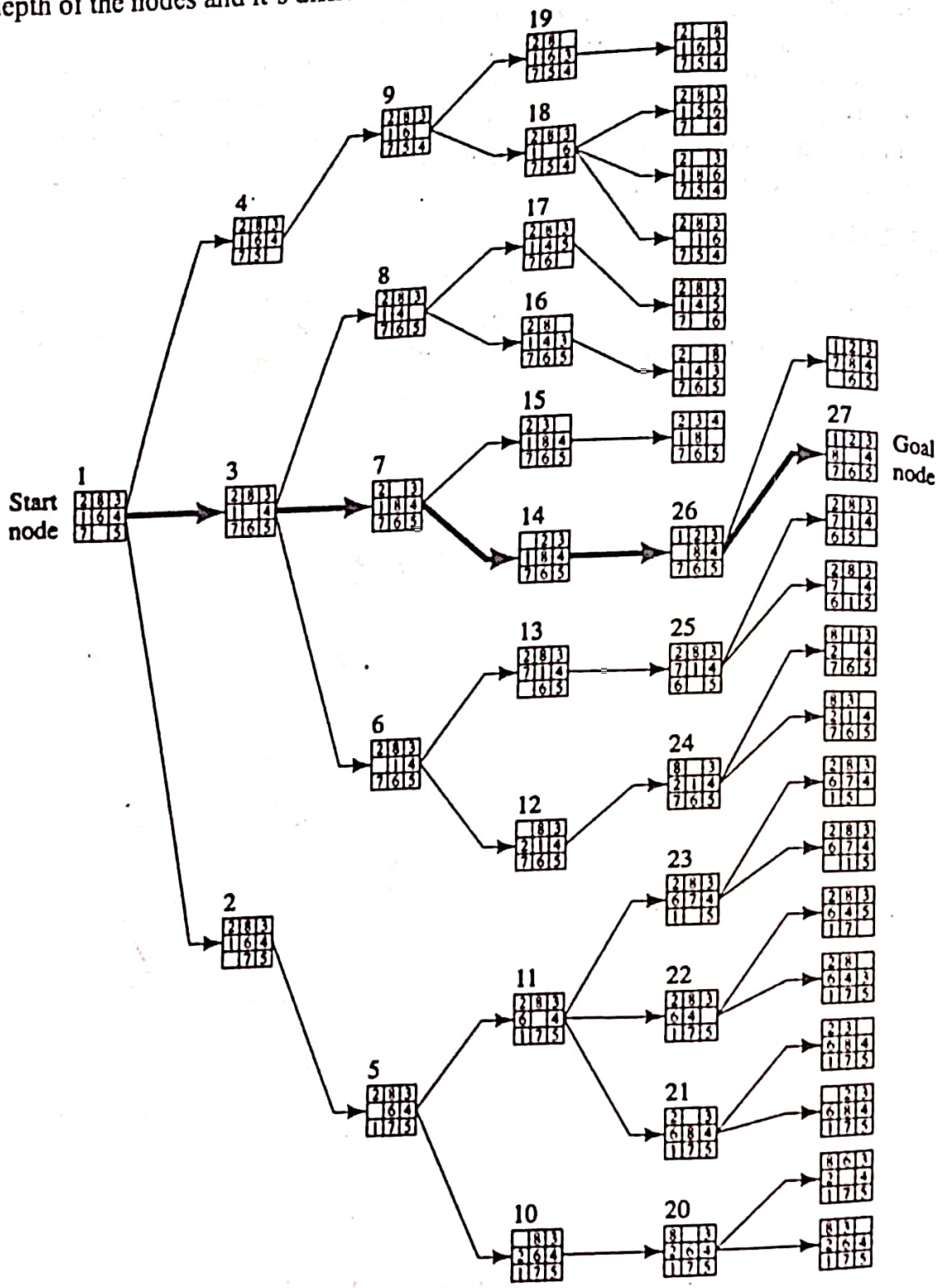
- **DFS for Binary search tree:**



DFS requires storing single path from initial to depth end node. DFS requires less memory to store the data. DFS may find the solution without examine the search space of

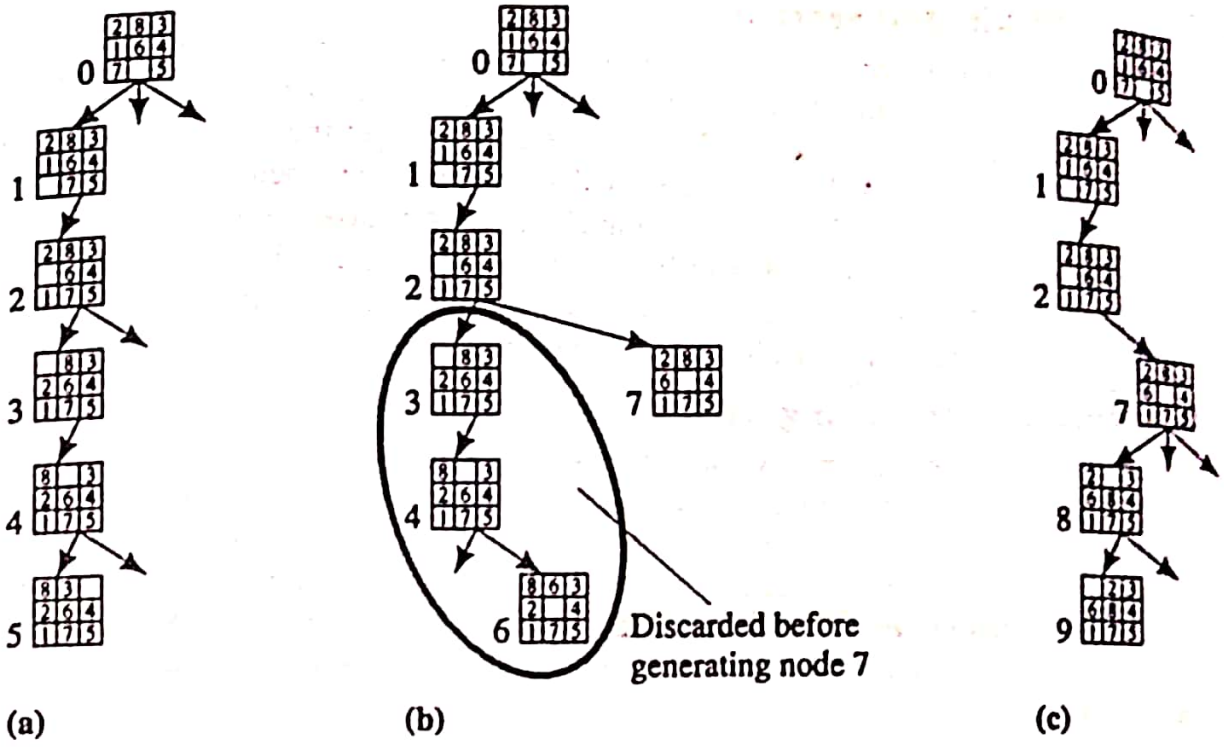
Scanned with CamScanner

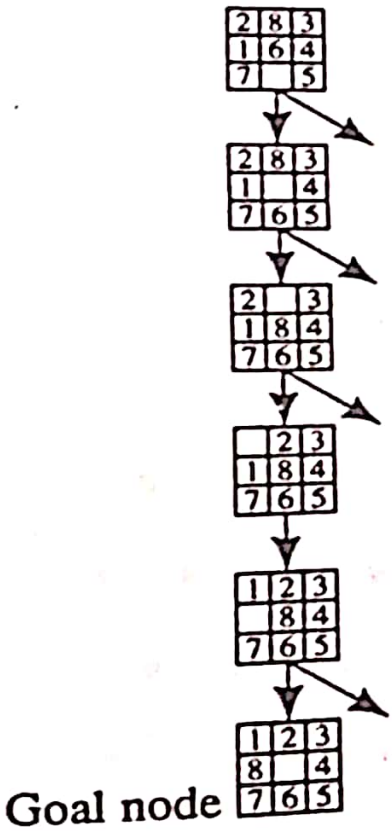
all. DFS is getting stop when one is formed. The only demerit of DFS is it may stick into the depth of the nodes and it's difficult to come back on the position (initial node).



BFS of Eight -Puzzle

Scanned with CamScanner





DFS of Eight Puzzles

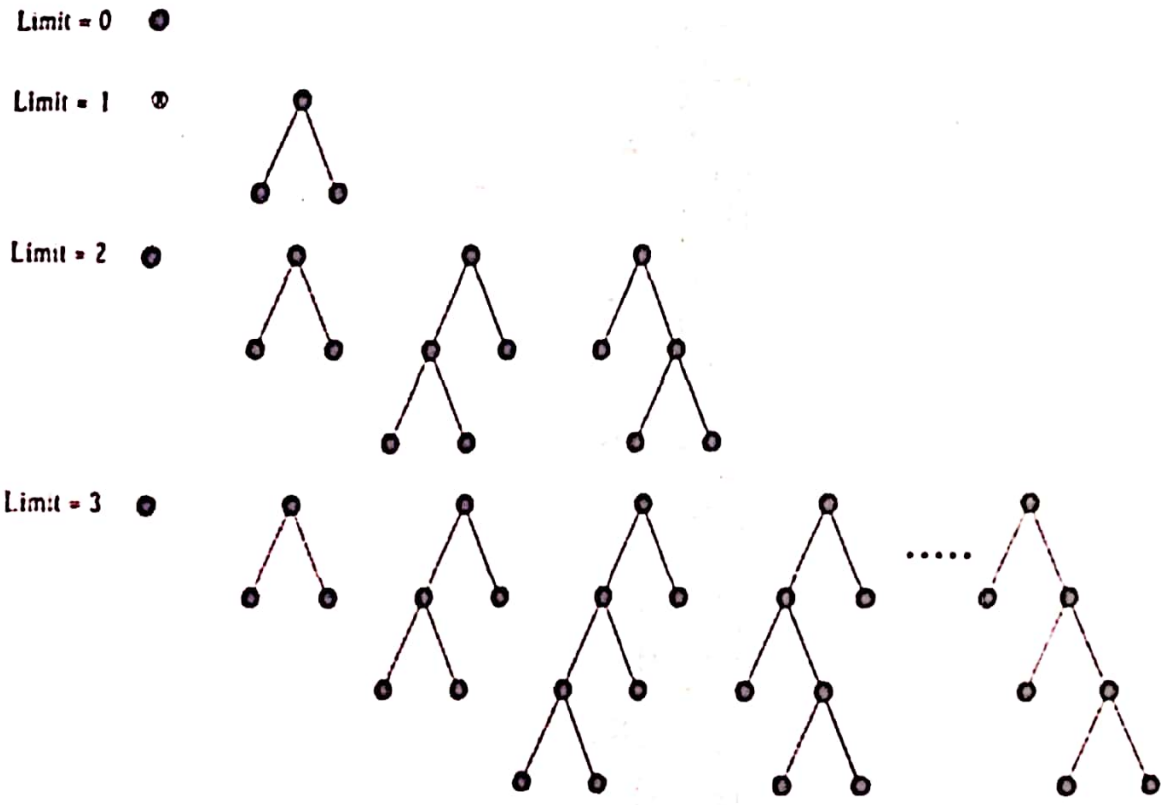
Scanned with CamScanner

2.5.3 Depth Limited search:

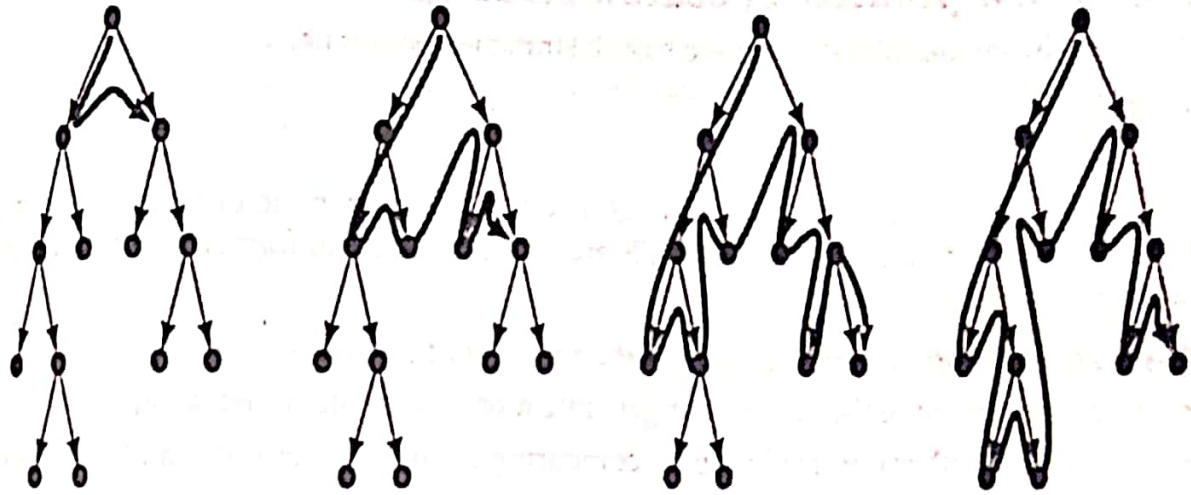
In depth limited search the search causes the minimum depth of the path on the nodes. For example: "Travelling Salesman Problem" where a salesman has to cross 20 cities to reach Udaipur from Delhi so the limit of the cities we get the length of 21 as the total cities is 22 including Delhi – 20 cities – Udaipur. If we implement the depth limit search and if salesman was in Delhi and to travel 21 cities it take the path of 21 steps, then generate a new state in city after Delhi (say it as Agra) or city B with the path of length that is one greater. With the help of this search we can find the solution that exists but still we cannot find the shortest path. The time and space taken of DLS is same as DFS.

2.5.4 Iterative Deepening DFS:

In DLS we choose the limit by own. In travelling salesman problem we random take the limit of city B but to reach Udaipur from many cities the distance is low as this number is known as diameter of state space which gives us better depth limit and more efficient DLS. Iterative deepening search try all the possible depth limits or we can say that iterative deepening search is a combination benefits of BFS and DFS. Sometimes in IDS so many stats are expanded more than two time which can be the waste of time and memory.



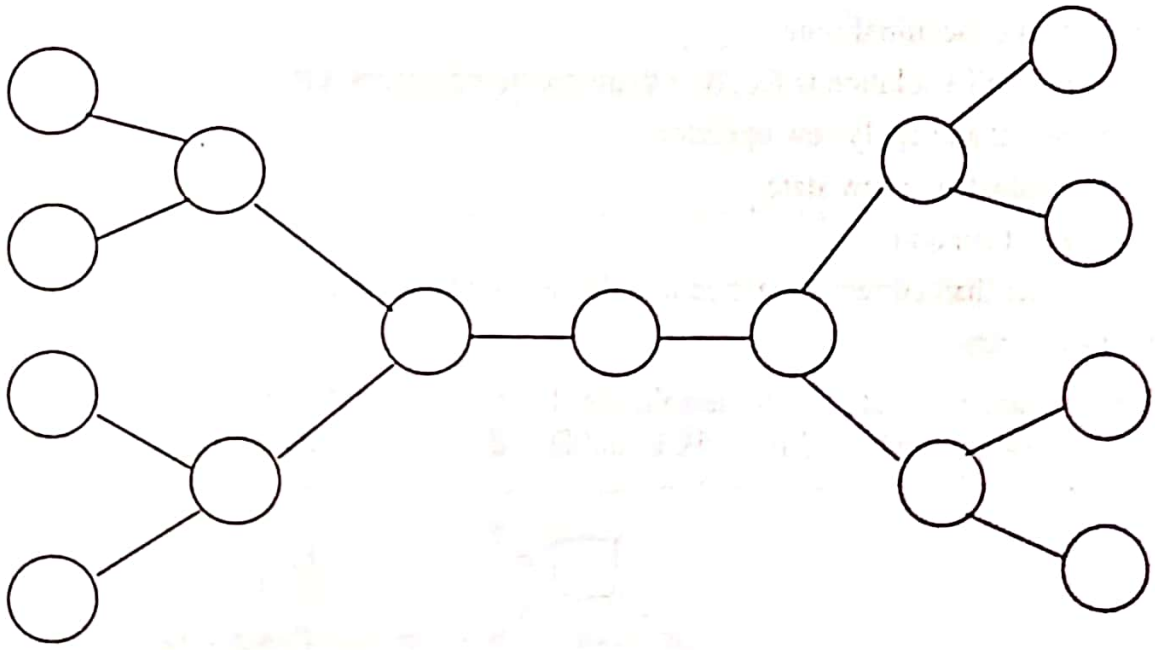
Scanned with CamScanner



Stages in Iterative –Deepening Search

2.5.5 Bidirectional Search:

The idea behind the bidirectional search is to search from both forward and backward, it can be search from initial state and also from end goal state.



When both the searching met in the middle of the problem it makes a big difference. It has to generate both successor and predecessor nodes to search from both the sides, theoretically the address of bidirectional search can be succeeded. Both the predecessor and successor are identical. Bidirectional search can reduce time complexity.

Scanned with CamScanner

2.6 Informed (Heuristic) search strategies:

We have already discussed two basic search strategies before i.e.

- BFS
- DFS

The searching proceeds preferably through nodes that are heuristic or informed are called heuristic search strategies. BFS and DFS are non heuristic (uniformed) search strategies. We here discuss following:

Generate and test: Generate and test is the simplest of all approaches.

- Generate a point in the problem or generate a path from the initial state.
- Then test whether it is a solution by comparing from the other states and goal state.
- If the solution is found then go to end point or else return to initial state.

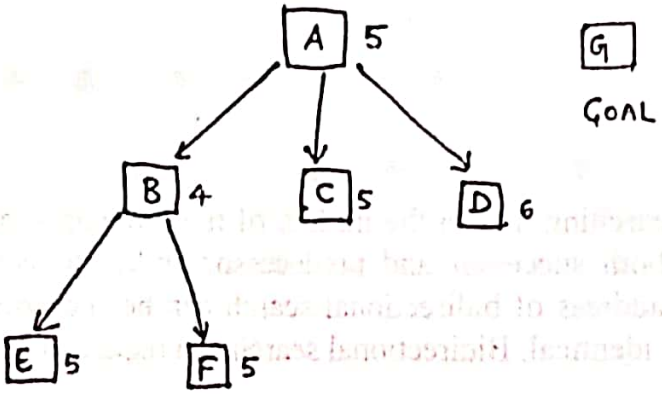
So if the possibility of solution is done systematically then this strategy can find many solutions. The best ways to apply generate and test strategy as a DFS with the backtracking method.

➢ **Hill Climbing Algorithm:**

This is the local search algorithm which has no backtracking and has a greedy approach.

1) Evaluate the initial state.
2) Loop until a solution is found or there are no operators left.
   - Select and apply new operator
   - Evaluate the new state
3) If goal then quit
   If better than current state then it is the new current state
4) Else return

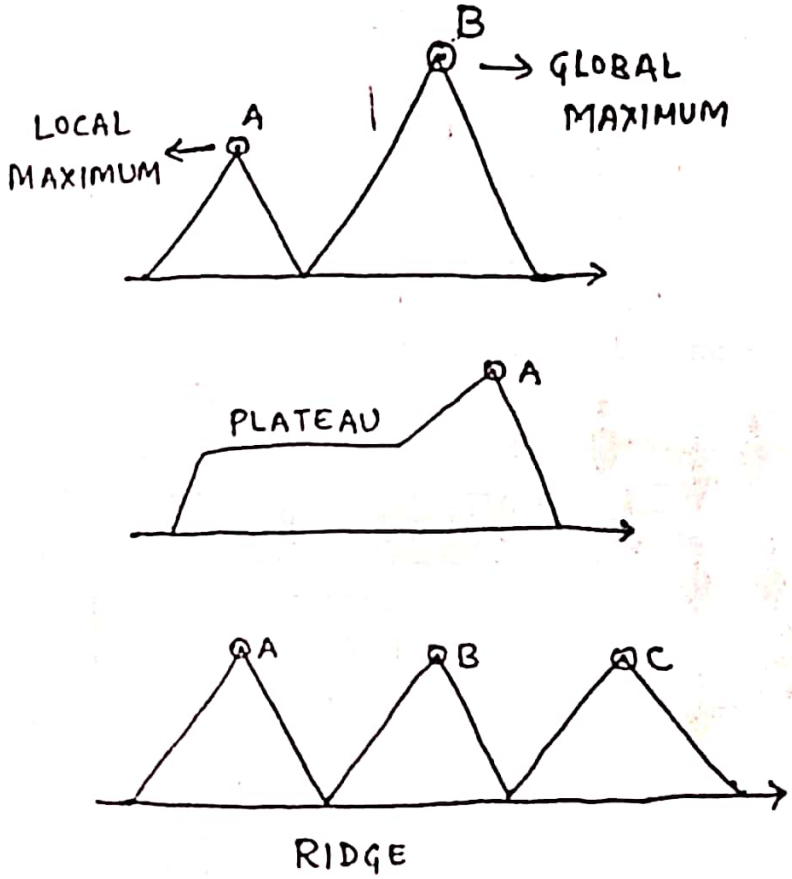
Get the maximum path to get the solution. Let us assume A is the initial state and 5 is its heuristic value. We have 3 paths B, C and D and because of its heuristic value best path is from B.



Scanned with CamScanner

➢ **Limitations:**

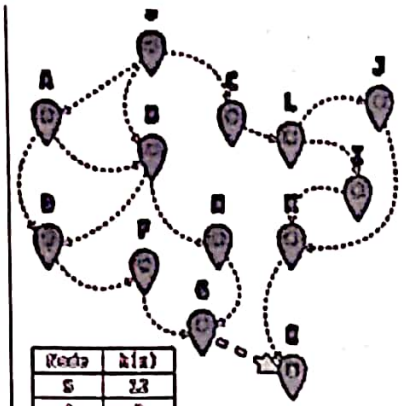
1. Local maximum- Let us suppose the algorithm get on point A, it assume that A is the destination point but when it goes down then it get point B which is the exact goal we have to achieve and it is the best solution.
2. Plateau- Plateau means a point where the values of all successors have same value but after sometime it become higher but for particular time algorithm get same value and it stopped and react that this is the last value which achieved but actually this is not the goal state.
3. Ridge- Ridge means at all the point where all the successors have same values and the first point searching gets stopped.



2.6.1 Concept of greedy Best first search:

In BFS the best part is selected with the best evaluation. In the greedy search according to the smallest heuristic value the number of steps taken to reach the goal node is heuristic value and the path which has small heuristic value picks the best node out of them as the next node to expand according some rules. We can say that h(n) is the cost value of heuristic.

Scanned with CamScanner



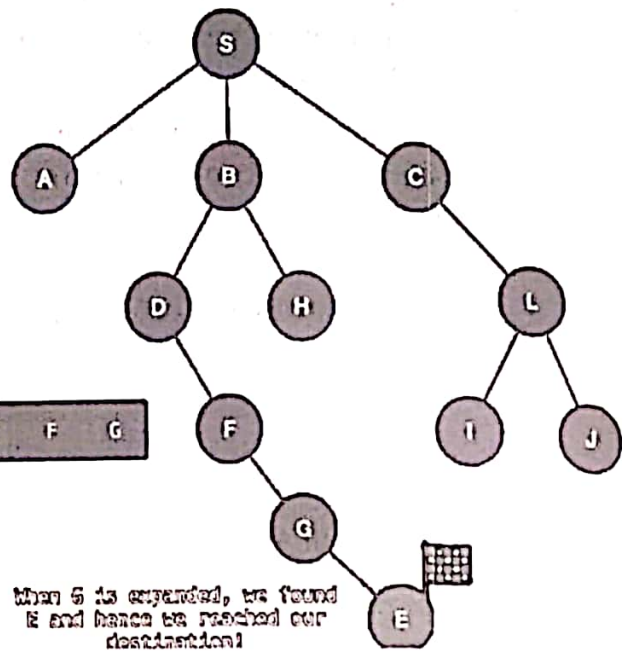
Node	h(n)
S	12
A	9
B	8
C	7
D	6
H	7
F	4
G	3
I	9
J	9
L	6
E	0

Nodes	S	C	L	B	D	F	G

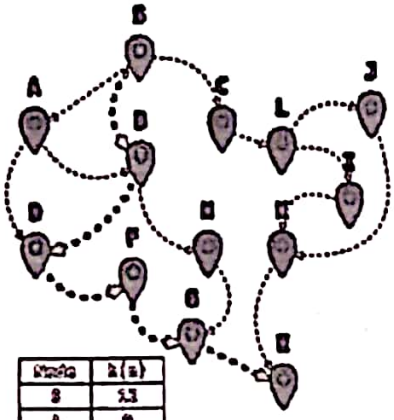
Closed List

Nodes	A	I	J	H
f(n)	9	9	9	7

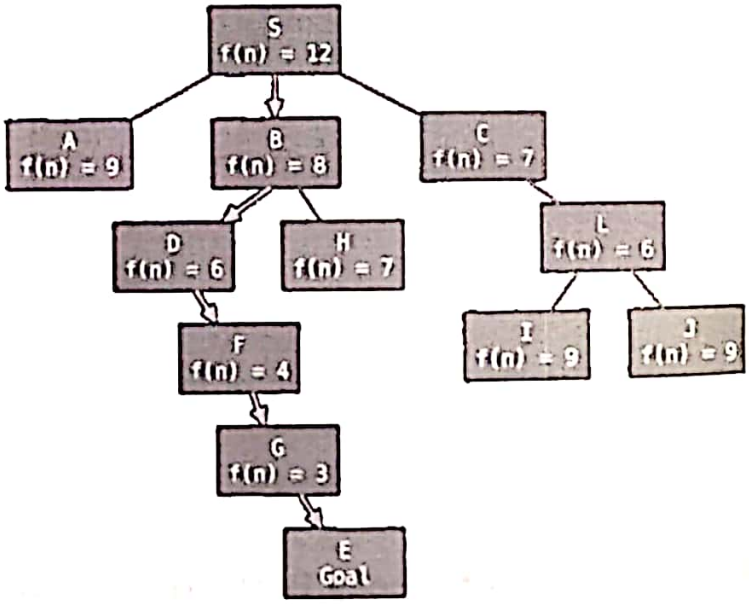
Open List



When G is expanded, we found E and hence we reached our destination!



Node	h(n)
S	12
A	9
B	8
C	7
D	6
H	7
F	4
G	3
I	9
J	9
L	6
E	0



Final tree with solution path highlighted in Green

Scanned with CamScanner

Algorithm

- Step 1: Place the starting node into the OPEN list.
- Step 2: If the OPEN list is empty, Stop and return failure.
- Step 3: Remove the node n from the OPEN list which has the lowest value of h(n), and place it in the CLOSED list.
- Step 4: Expand the node n and generate the successors of node n.
- Step 5: Check each successor of node n and find whether any node is a goal node or not. If any successor node is goal node, then return success and terminate the search, else proceed to Step 6.
- Step 6: For each successor node, algorithm checks for evaluation function f(n), and then check if the node has been in either OPEN or CLOSED list. If the node has not been in both list, then add it to the OPEN list.
- Step 7: Return to Step 2.

Greedy Best first search is useful in Robotics which can be commercial or domestic use. It also useful in Game zones and route planning.

2.6.2 A* search: minimizing the total estimated solution cost:

It is the best known form of Best first search. It avoids expanding of paths that are already expensive and it expandson most promising paths first.

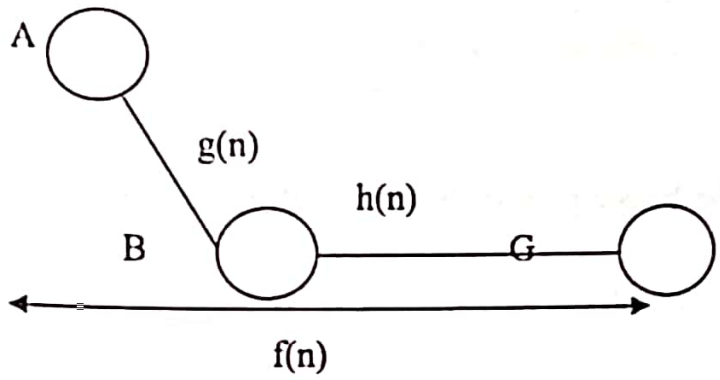
Thus,

$$F(n) = g(n) + h(n)$$

F(n)= estimated total path to reach goal

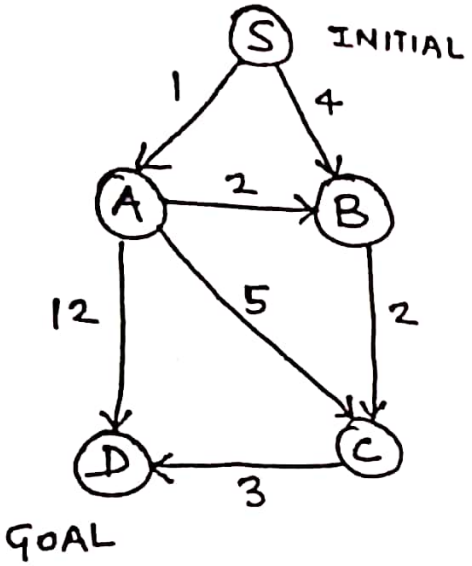
g(n)= cost to reach the node

h(n)= heuristic value



Scanned with CamScanner

Let us learn with the following example-



Heuristic Value	
S	7
A	6
B	2
C	1
D	0

S →

f(n) = g(n)+h(n)

= 0+7

=7 & so on..

Similarly we are doing on each state calculate the heuristic value for every path and select the smallest value.

S → A

S → B

S → B → C

S →A → B

S →A → C

S →A → D

S →A → B →C

S →A → B →C →D

S →B → C →D

S →A → C →D

After searching the entire path the below path is the best path to find the solution of a problem.

S →A → B →C →D This is the best path.

Scanned with CamScanner

2.7 Case Study: Applications of AI in transportation:

❖ Outline:

Artificial Intelligence is changing the transportation field. It helps cars, trucks, trains, ships and airplanes to function automatically to make traffic smooth, safe, clean, smart and more efficient. Due to increase of population, the evolution of transportation increases and it increase the complexity to manage and to analyze the generated data in transportation. To simplify and to maintain the complexity of such requirements AI is being used. AI-based techniques are applied in different form and in different areas whether it is road, aviation, waters or railways. AI helps all transportation safer and smarter. It also reduces human errors which involve traffic accidents and many more. The use of AI in transportation is also known as **intelligent transportation systems (ITS)** which include simplification of huge amount of data generated during transportation whether it is from vehicles as well as drivers.

❖ Method:

Transportation is the basic living standard for the present life. About 305 of the humans spent approx 2 hours of time in a day in travelling. In AI there are many methods to improve the use of transportation system. Particularly the transportation use AI as self control vehicles, predictions of path, and predictions of traffics. It can be divided as-

- Vehicle control system
- Traffic control system
- Traffic prediction system
- Accident prediction system
- Road safety system

The above AI methods decrease the level of accidents and injuries, because of the auto drivers, vehicles controlled so that less accident happens. Some methods are auto driving auto braking system, auto management, and emission control system. Predictions of traffic also reduce the traffic jams and accidents. Road safety systems help to avoid and prevent the possible accidents. Many transports are based on sensors such as GPS, cameras, radar in combination with the devices which transform an input signal into motion. Control systems and software are some of the methods which takeover the drivers functions some are parking the vehicle, Reversing, Driving, Stay, Stop which replaces the human drivers.

❖ Techniques:

AI is the vast area and there are many techniques which are applying at present and also in the past in all the area. In transportation area we grouped some of them are:

1. ANN as Artificial neural network
2. GAS as Genetic algorithm system

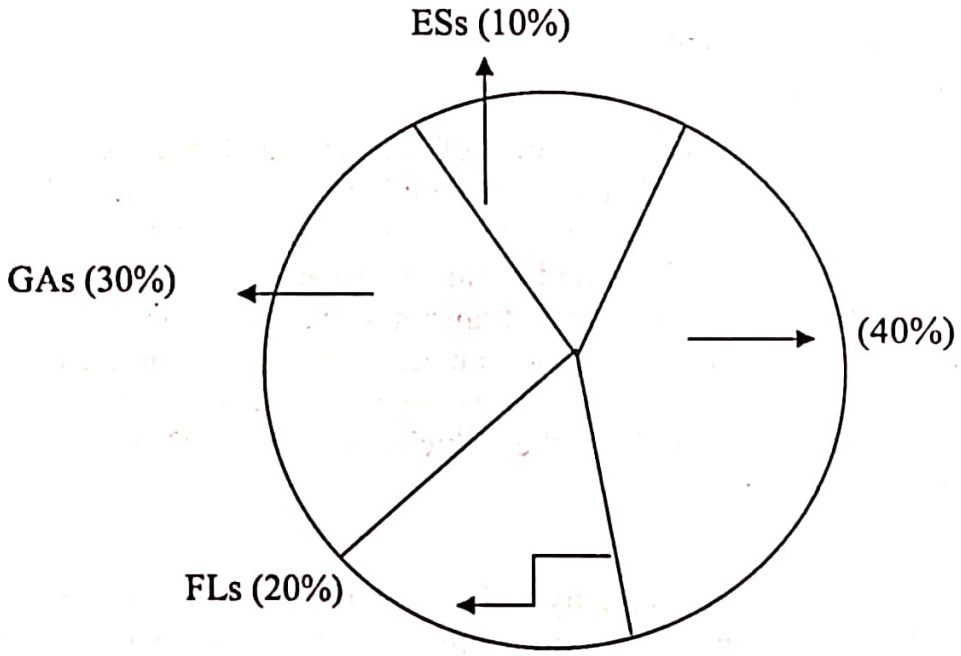
Scanned with CamScanner

3. FLS as Fuzzy logic system
4. ES as Expert system

Use of AI techniques in different transportation areas

Areas	AI Techniques			
	ANN	GAs	FL	ESs
Vehicle control system	5	8	5	2
Traffic control system	11	4	7	1
Road safety and accident prediction	5	3	6	3

It analyze that artificial neural network technique is the most used techniques among all. If we take an example of roads and traffics ANN and FLs are used to monitoring and to find the stability of vehicles in intelligent system and to avoided traffic jams and predictions of vehicles GAs implemented.

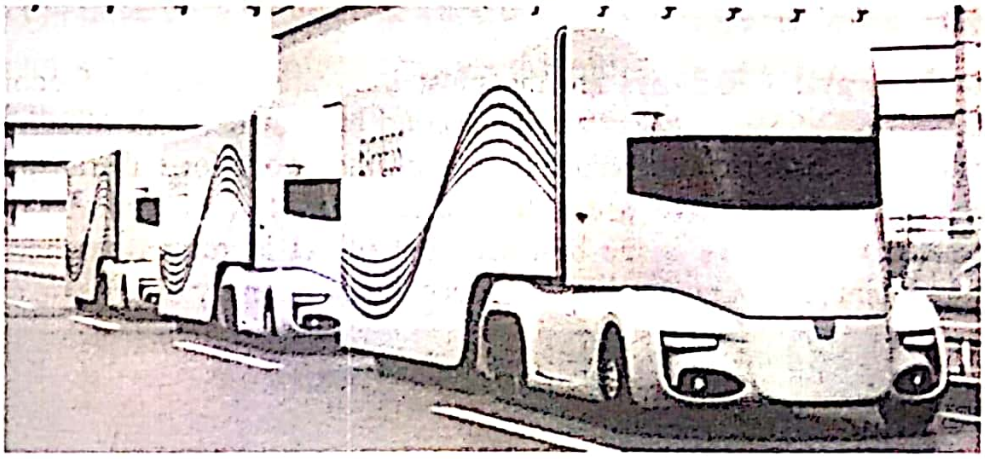


Scanned with CamScanner

Areas:

As we already mention with illustrations the areas where AI is used, here we discuss in details.

- **AI in Roads: - Illustration of truck platooning.**



AI makes truck platooning, let 4 vehicles are inline the driver of the first vehicle can control other three trucks or allowing them to let go of the steering wheel and let the truck steer by own or it is also called as automatic driving system. The distance between the vehicles is about to 10-12 meters. This method is very helpful which cause less congestion. Another platform where AI is used in uber drivers were the riders and drivers can recognizes the identity and also the route optimization.

- **AI In Aviation: - Illustration of automated pilots in wars.**

Scanned with CamScanner

It's about 20 million of flights in 2018. AI is not new in the aviation industry by automating driving, brake, traffic predictions. AI helps aviation systems for safe and secure. AI makes aviation business helpful with machine learning it predict the data what the customer want through google or social media and adjust the content according to customers. Some highly uses cases are fuelling, catering, loading-unloading passengers, baggage control, emails and security.

- **AI in Railways: -** Railways are the most innovative sectors of economy and major industrial revolution. Railways are the second biggest network after web. AI improves infrastructure, manufacture, maintenance and other work for rail transports. It helps with lower cost, safe, effective and a big competition to other modes of transportation. Automated metro trains is the best use case of AI in railways where the source, the distance, point of stay, minute of stay reduce the complexity of this use case.

Scanned with CamScanner

Exercises

❖ Answer the following Questions in brief.

1. Explain the difference between BFS and DFS?
2. Explain Informed and Uninformed strategies with a Toy problem
3. What is the concept of Greedy BFS and its algorithm? Explain with real time example?
4. Explain the Applications of Artificial Intelligence in –
   - Gaming
   - Social Media

❖ Multiple Choice Questions:

1. What is the main task of a problem-solving agent?
   A. Solve the given problem and reach to goal
   B. To find out which sequence of action will get it to the goal state
   C. Both A and B
   D. None of the Above

2. What is Initial state + Goal state in Search Terminology?
   A. Problem Space B. Problem Instance
   C. Problem Space Graph D. Admissibility

3. Depth-First Search is implemented in recursion with _____ data structure.
   A. LIFO B. LILO C. FIFO D. FILO

4. How many types are available in uninformed search method?
   A. 2 B. 3 C. 4 D. 5

5. How many types of informed search method are in artificial intelligence?
   A. 2 B. 3 C. 4 D. 5

6. 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

7. State whether the following statements about the state space are True.
   i) A state-space forms a graph in which the nodes are states and the arch between nodes are actions.
   ii) In state space, a path is a sequence of states connected by a sequence of actions.
   A. i-only B. ii-only C. Both i and ii D. None of the above

Scanned with CamScanner

8. The Set of actions for a problem in state space is formulated by which one of the following?
   A. Successor function, which takes current action and returns next immediate state
   B. Initial state C. Intermediate states D. None of these
9. Which of the following is a touring problem in which each city must be visited exactly once? The purpose is to search for the shortest tour among all the tours.
   A. Searching the shortest path between a source and a destination
   B. Depth-first search traversal on a given map represented as a graph
   C. Map coloring problem D. Travelling Salesman problem
10. Which of the following searching technique takes less memory?
   A. Optimal search B. Breadth-First Search
   C. Linear Search D. Depth-First Search

Answer:

1. **Ans : C**
   Explanation: The problem-solving agents are one of the goal-based agents.
2. **Ans : B**
   Explanation: Problem Instance : It is Initial state + Goal state.
3. **Ans : A**
   Explanation: Depth-First Search implemented in recursion with LIFO stack data structure.
4. **Ans : D**
   Explanation: The five types of uninformed search method are Breadth-first, Uniform-cost, Depth-first, Depth-limited and Bidirectional search.
5. **Ans : C**
   Explanation: The four types of informed search method are best-first search, Greedy best-first search, A* search and memory bounded heuristic search.
6. **Ans : C**
   Explanation: Sometimes minimum heuristics can be used, sometimes maximum heuristics function can be used. It depends upon the application on which the algorithm is applied.
7. **Ans: C. Both i and ii**
8. **Ans: A** 9. **Ans: D** 10. **Ans: D**

Scanned with CamScanner

UNIT-3

Natural Language Processing

- Language Models
  - N-gram character models
  - N-gram word models
- Text classification
  - Classification by data compression
- Information retrieval
  - The page rank algorithm
  - The HITS algorithm
- Information extraction
  - Finite state automata for information extraction
  - Probabilistic model for information extraction
- Examples: Applications of Natural Language Processing.
- Case Study: Automated Voice Assistants, Chat bots.

Scanned with CamScanner

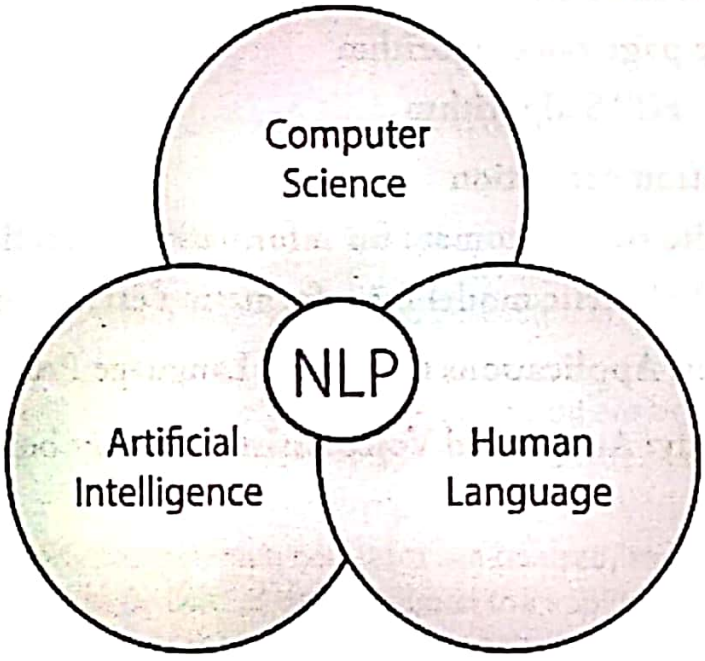
Unit –3 Natural Language Processing

❖ Introduction:

NLP: the ability to understand and process human languages. It is important in order to fill the gap in communication between humans and machines.

Natural Language Processing (NLP), by definition, is a method that enables the communication of humans with computers or rather a computer program by using human languages, referred to as natural languages, like English. These include both text and speech input. It helps computers to understand and interpret the languages and reply validly in a valid manner.

It is a part of **Computer Science, Human language,** and **Artificial Intelligence**. It is the technology that is used by machines to understand, analyze, manipulate, and interpret human's languages. It helps developers to organize knowledge for performing tasks such as **translation, automatic summarization, Named Entity Recognition (NER), speech recognition, relationship extraction,** and **topic segmentation.**It is used to create automated software that helps understand human spoken languages to extract useful information it gets within the style of audio. Techniques in NLP allow computer systems to process and interpret data in the form of natural languages.



Natural language processing (NLP) drawbacksare often divided into two tasks:

- Processing written text, using lexical, syntactic and semantic information of the language as well as the required real world information.

Scanned with CamScanner

- Processing spoken language, using all the information required higher than and extra knowledge about phonology as well as enough added information to handle the additional ambiguities that arise in speech.
- NLP is one of the major and most addressed parts of Artificial Intelligence. It is very commonly used in our day-to-day lives, in applications like Google Assistant, Siri, Google Translate, Alexa, etc. NLP includes dividing the input into smaller items and playing tasks to understand the connection between them and then produce significant output.

3.1 Language Models:

Formal languages, such as the programming languages Java or Python have exactly defined language models. A **language** can be defined as a collection of strings; "print (2 + 2)" is a legal program in the language Python, whereas "2) + (2 print" is not. Since there are an infinite number of legal programs, they cannot be enumerated; instead they are such that by a collection of rules called a **grammar**. Formal languages also have rules that define the meaning or **semantics** of a program; as an example, the rules say that the "meaning" of "2 + 2" is 4, and the meaning of "1/0" is that an error is signaled.

Natural languages are also **ambiguous**. We cannot speak of a single meaning for a sentence, but rather of a probability distribution over possible meanings.

3.1.1 N-gram character models:

An n-gram model is **a technique of counting sequences of characters or words that allows us to support rich pattern discovery in text.** In other words, it tries to capture patterns of sequences (characters or words next to each other) while being sensitive to contextual relations (characters or words near each other).

Ultimately, a written text is consist of **characters**—letters, digits, punctuation, and spaces in English (and additional exotic characters in some other languages). Thus, one of the simplest language models is a probability distribution over sequences of characters. We have tendency to write P (c1:N) for the probability of a sequence of N characters, c1 through cN.

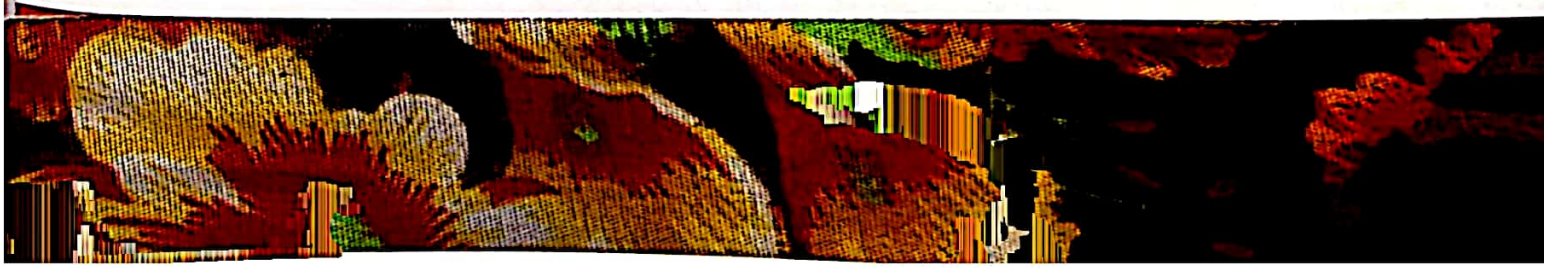
A sequence of written symbols of length n is called an n-gram, with special case "unigram" for 1-gram, "bigram" for 2-gram, and "trigram" for 3-gram. A model of the probability distribution of n-letter sequences is thus called an **n-gram model.**

An n-gram model is defined as a **Markov chain** of order n − 1.

$$P(c_i \mid c_{1:i-1}) = P(c_i \mid c_{i-2:i-1})$$

We can define the probability of a sequence of characters P (c1: N) under the trigram model by first factoring with the chain rule and then using the Markov assumption:

$$P(c_{1:N}) = \prod_{i=1}^{N} P(c_i \mid c_{1:i-1}) = \prod_{i=1}^{N} P(c_i \mid c_{i-2:i-1}) .$$



Scanned with CamScanner

For a trigram character model in a language with 100 characters, **P** (Ci|Ci−2:i−1) has a million entries, and can be accurately estimated by counting character sequences in a body of text of 10 million characters or more. We call a body of text a **corpus.**

What can we do with n-gram character models? One task for which they are well suited is **language identification**: given a text, determine what natural language it is written in. This is a relatively easy task; even with short texts such as "Hello, world" or "Wiegehtesdir," it is easy to identify the first as English and the second as German. Computer systems identify languages with greater than 99% accuracy; sometimes, closely related languages, such as Swedish and Norwegian are confused.

Other tasks for character models include spelling correction, genre classification, and named-entity recognition. **Genre classification** means deciding if a text is a news story, alegal document, a scientific article, etc. While many features help make this classification,counts of punctuation and other character n-gram features go a long way. Named-entity recognition is the task of finding names of things in a document and deciding what class they belong to.

For example, in the text "Mr. Sopersteen was prescribedaciphex," we should recognize that "Mr. Sopersteen" is the name of a person and "aciphex" isthe name of a drug. Character-level models are good for this task because they can associatethe character sequence "ex" ("ex" followed by a space) with a drug name and "Steen" witha person name, and thereby identify words that they have never seen before.

3.1.2 N-gram word models:

When we have a tendency to analyze a sentence one word at a time, then it is referred to as a unigram. The sentence parsed two words at a time isa bigram.

When the sentence is parsed three words at a time, then it is a trigram. Similarly, n-gram refers to the parsing of *n* words at a time.

Example: To understand unigrams, bigrams, and trigrams, you can refer to the below diagram:

Scanned with CamScanner

Therefore, parsing allows machines to understand the individual meaning of a word in a sentence. Also, this type of parsing helps predict the next word and correct spelling errors.

Consider two sentences: "There was heavy rain" vs. "There was heavy flood". From experience, we all know that the previous sentence sounds better. An N-gram model will tell us that "heavy rain" occurs much more often than "heavy flood" within the training corpus. Thus, the first sentence is more probable and can be elected by the model.

A model that simply depends on however a word occurs without looking at previous words is called **unigram**. If a model considers only the previous word to predict the current word, then it's called **bigram**. If two previous words are considered, then it's a **trigram** model.

An n-gram model for the above example would calculate the following probability:

P('There was heavy rain') = P('There', 'was', 'heavy', 'rain') =
P('There')P('was'|'There')P('heavy'|'There was')P('rain'|'There was heavy')

Since it's impractical to calculate these conditional probabilities, using *Markov assumption*, we approximate this to a bigram model:

P('There was heavy rain') ~ P('There')P('was'|'There')P('heavy'|'was')P('rain'|'heavy')

Now we turn to n-gram models over words instead of characters. All the same mechanismapplies equally to word and character models. The main difference is that the **vocabulary**the set of symbols that make up the corpus and the modelis larger. There are only about100 characters in most languages, and sometimes we build character models that are evenmore restrictive, for example by treating "A" and "a" as the same symbol or by treating allpunctuation as the same symbol. But with word models we have at least tens of thousands ofsymbols, and sometimes millions. The wide range is because it is not clear what constitutes aword. In English a sequence of letters enclosed by spaces is a word, but in some languages,like Chinese, words are not separated by spaces, and even in English many decisions must bemade to have a clear policy on word boundaries: how many words are in "ne'er-do-well"? Orin "(Tel:1-800-960-5660x123)"?

3.2 Text classification:

Text classification is the process of classifying documents into predefined categories based on their content. It is the machine controlled (automated) assignment of natural language texts to predefinedclasses (or categories). Text classification is the primary requirement of text retrieval systems thatretrieve texts in response to a user query, and text understanding systems, which transform text in some way such as producing summaries, questions-answers or extract data.

We now consider in depth the task of **text classification**, also known as **categorization:** given a text of some kind, decide that of a predefined set of categories it belongs to.

Scanned with CamScanner

Text classification is that the method of (process of) classifying documents into predefined categories by their content.

Language identification and genre classification are examples of text classification, as issentiment analysis(classifying a movie or product review as positive or negative) and **spam detection** (classifyingan email message as spam or not-spam). Since "not-spam" is awkward, researchers havecontained the term **ham** for not-spam. We can treat spam detection as a problem in supervisedlearning. A training set is readily available: the positive (spam) examples are in my spamfolder, the negative (ham) examples are in my inbox. Here is an excerpt:

Spam: Wholesale Fashion Watches -57% today. Designer watches for cheap ...

Spam: You can buy ViagraFrS1.85 All Medications at unbeatable prices! ...

Spam: WE CAN TREAT ANYTHING YOU SUFFER FROM JUST TRUST US...

Spam: Sta.rt earn*ing the salary yo,u d-eserve by o'btaining the prope,rcrede'ntials!

Ham: The practical significance of hyper tree width in identifying more...

Ham: Abstract: We will motivate the problem of social identity clustering: ...

Ham: Good to see you my friend. Hey Peter, It was good to hear from you. ...

Ham: PDS implies convexity of the resulting optimization problem (Kernel Ridge ...

From this section(part) we can start to get an idea of what might be good features to include inthe supervised learning model. Word n-grams such as "for cheap" and "You can buy" seemto be indicators of spam (although they would have a nonzero probability in ham as well).

Character-level features also seem important: spam is more likely to be all uppercase and tohave punctuation embedded in words. Apparently the spammers thought that the word bigram"you deserve" would be too indicative of spam, and thus wrote "yo,u d-eserve" instead. Acharacter model should detect this. We could either create a full character n-gram modelof spam and ham, or we could handcraft features such as "number of punctuation marksembedded in words."

Note that we have two complementary ways in which of talking about classification. In thelanguage-modeling approach, we define one n-gram language model for P(Message | spam)by training on the spam folder, and one model for P(Message | ham) by training on the inbox.

Then we can classify a new message with an application of Bayes' rule:

$$\underset{c \in \{spam, ham\}}{\operatorname{argmax}} P(c \mid message) = \underset{c \in \{spam, ham\}}{\operatorname{argmax}} P(message \mid c)\, P(c).$$

Scanned with CamScanner

Where,P(c) is estimated just by counting the total number of spam and ham messages. Thisapproach works well for spam detection,just as it did for language identification.

The data can accurately determine if it is good or not. It is necessary to constantlyupdate features, because spam detection is an **adversarial task**; the spammers modify theirspam in response to the spam detector's changes. It can be expensive to run algorithms on a very large feature vector, so often a processof **feature selection** is used to keep only the features that best discriminate between spam andham.

3.2.1 Classification by data compression:

Another way to think about classification is as a problem in **data compression**. A losslessCompression algorithm takes a sequence of symbols, detects repeated patterns in it, and writesa description of the sequence that is lot of compact than the original. For example, the text"0.142857142857142857" might be compressed to "0. [142857]*3." Compression algorithmswork by dictionaries of subsequences of the text, and then referring to entries within thedictionary. The example here had only one dictionary entry, "142857."

In effect, compression algorithms are creating a language model. To do classificationby compression, first collect all the spam training messages and compress them as aunit. We do the same for the ham. Then when given a new message to classify, we appendit to the spam messages and compress the result. We also append it to the ham and compressthat. Whichever class compresses betteradds the fewer number of additional bytes for thenewmessageis the predicted class. The idea is that a spam message will tend to share dictionary entries with different spam messages and therefore can compress better when appended toa collection that already contains the spam dictionary.

Experiments with compression-based classification on number of the standard corpora fortext classification the 20-Newsgroups data set, the Reuters-10 Corpora, the Industry Sectorcorporaindicate that whereas running off-the-shelf compression algorithms like gzip, RAR can be quite slow; their accuracy is comparable to traditional classification algorithms.

This is interesting in its own right, and also serves to point out that there's promisefor algorithms that use character n-grams directly with no preprocessing of the text or featureselection: they appear to be capturing some real patterns.

3.3 Information retrieval :

Information retrieval is concerned with representing, searching, and manipulating large collections of electronic text and alternative human-language information.

Information retrieval systems use a very simple language model based on **bags ofwords**, yet still manage to perform well in terms of **recall** and **precision**(exactness)on very largecorpora of text. On internet corpora, link-analysis algorithms improve performance.

Scanned with CamScanner

Information retrieval is the task of finding documents that are relevant to a user's need forinformation. The well-known examples of information retrieval systems are search enginesontheWorldWideWeb. A Web user can type a query such as: AI bookinto a search engineand see a list of relevant pages. In this section, we will see how such systems are built. Aninformation retrieval system can be characterized by

1. **A corpus of documents.** Every system should decide what it needs to treat as a document: A paragraph, a page, or a multipage text.
2. **Queries posed in a query language.** A query specifies what the user needs to understand.

The query language will be just a list of words, such as [AI book]; or it can specifya phrase of words that has to be adjacent, as in ["AI book"]; it will contain Booleanoperators as in [AI AND book]; it can include non-Boolean operators such as [AI NEAR book] or [AI book site: www.aaai.org].

3. **A result set.** This is the subset of documents that the IR system judges to be **relevant** tothe query. By *relevant*, we tend to mean likely to be of use to the person who posed the query,for the particular information need expressed in the query.
4. **A presentation of the result set.** This will be as simple as a ranked list of documenttitles or as advanced as a rotating color map of the result set projected onto a threedimensionalspace, rendered as a two-dimensional display.

3.3.1 The page rank algorithm:

It is a scoring measure based only on the link structure of web pages or website. A web page is important if it is pointed to by other important web pages. Our first technique for link analysis assigns to every node in the web graph a numerical score between 0 and 1 known as it page rank. Given a query, a web search engine computes a composite score for everyweb page content that combines or mixes hundreds of options like cos similarity and term proximity together with the Page rank score.

PageRank was one of the two original ideas that set Google's search except from other Web Search engines when it was introduced in 1997.

```
function HITS(query) returns pages with hub and authority numbers
  pages ← EXPAND-PAGES(RELEVANT-PAGES(query))
  for each p in pages do
    p.AUTHORITY ← 1
    p.HUB ← 1
  repeat until convergence do
    for each p in pages do
      p.AUTHORITY ← ∑i INLINKi(p).HUB
      p.HUB ← ∑i OUTLINKi(p).AUTHORITY
    NORMALIZE(pages)
  return pages
```


Scanned with CamScanner

The HITS algorithm for computing hubs and authorities with respect to aquery. RELEVANT-PAGES fetch the pages that match the query, and EXPAND-PAGES addsin each page that links to or is linked from one of the relevant pages. NORMALIZE divideseach page's score by the sum of the squares of all pages' scores (separately for both theauthority and hubs scores).

(The other innovation was the use of anchor text—the underlined text in a hyperlink—to index a page, even though the anchor text was ona*different*page than the one being indexed.) PageRank was invented to solve the problem ofthe tyranny of TF scores: if the query is [IBM], how do we make sure that IBM's home page,ibm.com, is the first result, even if another page mentions the term "IBM" more frequently?

The idea is that ibm.com has several in-links (links to the page), so it should be ranked higher:every in-link is a vote for the standard of the linked-to page. But if we only counted in links,then it would be possible for a Web spammer to create a network of pages and have all of them allpoint to a page of his selecting, increasing the score of that page.

Therefore, the PageRankalgorithm is designed to weight links from high-quality sites more heavily. What is a highqualitysite? Onethat is linked to by other high-quality sites. The definition is Algorithmic or recursive, butwe will see that the recursion bottoms out properly. The PageRank for a page p is defined as:

$$PR(p) = \frac{1-d}{N} + d\sum_i \frac{PR(in_i)}{C(in_i)},$$

Where PR (p) is the PageRank of page p, N is the total number of pages in the corpus, in_i is the pages that link in to p, and C (in_i) is the count of the total number of out-links on page in_i. The constant d is a damping factor. It can be understood through the **randomsurfer model**: imagine a Web surfer who starts at some random page and begins exploring.

With probability d (we'll assume d=0.85) the surfer clicks on one in all the links on the page (choosing uniformly among them), and with chance 1 − d she gets uninterested in the pageand restarts on a random page anywhere on the Web. The PageRank of page p is then theprobability that the random surfer will be at page p at any purpose in time. Page Rank can becomputed by an iterative procedure: starts with all pages having PR(p)=1, and iterate thealgorithm, updating ranks until they converge.

Thus PageRank could be global ranking of all web pages based on their locations in the web graphstructure. PageRank uses information that is external to the web pages – **backlinks**. Backlinks from important pages are moresignificant than backlinks from average pages.

Scanned with CamScanner

Advantages of Page Rank:

- **FightingSpam.**A page is important if the pages pointing to it are important. Since it is hard for Web page owner to add in-links into his/her page from different important pages, it is thus not easy to influence Page Rank.
- **Page Rank is a global measure and is query independent.** Page Rank values of all the pages are computed and saved off-line rather than at the query time.

3.3.2 The HITS algorithm:

Hyperlink Induced Topic Search (HITS) Algorithm is a Link Analysis Algorithm that rates Webpages,developed by Jon Kleinberg in 1999. This algorithm is used to the web link-structures to discover and rank theWebPages relevant for a particular search.

HITS use hubs and authorities to define a recursive relationship between webpages. To get knowledge of HITS Algorithm, we first have to get knowledge about Hubs and Authorities.

- Given a query to a Search Engine, the set of highly relevant web pages are called **Roots.** They arepotential **Authorities.**
- Pages that are not very relevant but point to pages in the Root are called **Hubs.** Thus, an Authority is apage that many hubs link and a Hub is a page that links tomany authorities.

Algorithm –

-> Let number of iterations be k.

->each node is assigned a Hub score = 1 and an Authority score = 1.

-> Repeat k times:

- ***Hub update:** Each node's Hub score = (Authority score of each node it points to).*
- ***Authority update:** Each node's Authority score = (Hub score of each node pointing to it).*
- *Normalize the scores by dividing each Hub score by square root of the sum of the squares of all Hub scores, and dividing each Authority score by square root of the sum of the squares of all Authority scores.*
- **Two sets of inter-related pages:**
  - ⇨ Hub Pages-good lists of links on a subject
  - ⇨ Authority pages-occur recurrently on good hubs for the subjects.

The HITS algorithm

$$H(x) \leftarrow -\sum a(y)$$

$$A(x) \leftarrow -\sum h(y)$$

The Hyperlink-Induced Topic Search algorithm, also known as "Hubs and Authorities" orHITS, is another influential link-analysis algorithm, HITS differs from Page Rank in

Scanned with CamScanner

several ways. First, it is a query-dependent measure: it rates pages with respectto a query. That means that it must be computed anew for each querya computationalburden that most search engines have elected not to take on. Given a query, HITS first findsa set of pages that are relevant to the query. It does that by intersecting hit lists of querywords, and then adding pages in the link neighborhood of these pages—pages that link to oris linked from one of the pages in the original relevant set.

Each page in this set is considered an **authority** on the query to the degree that otherpages in the relevant set point to it. A page is considered a **hub** to the degree that it pointstoother authoritative pages in the relevant set. Just as with PageRank, we don't want tomerely count the number of links; we want to give more value to the high-quality hubs andauthorities. Thus, as with PageRank, we iterate a process that updates the authority score ofa page to be the sum of the hub scores of the pages that point to it, and the hub score to bethe sum of the authority scores of the pages it points to. If we then normalize the scores andrepeat k times, the process will converge.

Both PageRank and HITS played important roles in developing our understanding ofWeb information retrieval. These algorithms and their extensions are used in ranking billionsof queries daily as search engines steadily develop better ways of extracting yet finer signalsof search relevance.

➢ **PageRank vs. HITS**

PageRank	HITS
Computed for all web pages stored prior to the query	Performed on the subset generated by each query
Computes authorities only	Computes authorities and hubs
Fast to compute	Easy to compute, realTime execution is hard.

➢ **Information Retrieval vs.Information Extraction:**

- **Information Retrieval**: Given a set of terms and a set of document terms select only the most relevant document (precision and preferably all the relevant ones (recall).
- **Information Extraction**: Extract from the text what the document means.

3.4 Information Extraction:

Information extraction (IE) is the task of automatically extracting structured information from unstructured and/or semi-structured machine-readable documents and other electronically represented sources. In most of the cases this activity concerns processing human language texts by means of natural language processing (NLP).

Scanned with CamScanner

- Information Extraction refers to the automatic extraction of structured information such as entities, relationships between entities, and attributes describing entities from unstructured sources.
- Identify phrases in language that refer to specific types of entities and relations in text
- Named entity recognition is the task of identifying names of people, places, organizations, etc. in text.
- Relation extraction identifies specific relations between entities. A man works for IBM -> PERSON works for ORGANIZATION

Information extraction is the process of acquiring knowledge by skimming a text and looking for occurrences of a particular class of object and for relationships among objects. A difficult task is to extract instances of addresses from Web pages, with database fields for street, city, state, and zip code.

3.4.1 Finite state automata for information extraction:

The simplest type of information extraction system is an **attribute-based extraction** system that assumes that the entire text refers to a single object and the task is to extract attributes of that object.

For example,the problem of extracting from the text "IBM ThinkBook970.Our price: Rs.399.00". the set of attributes

```
{
        Manufacturer=IBM,
        Model=ThinkBook970,
        Price=Rs.399.00
}
```

We can address this problem by defining a **template** (also known as **a pattern**) for each attribute we would like to extract. The template is defined by a finite state automaton, the simplest example of which is the **regular expression**, or **regex**.

Regular expressions are used in UNIX commands such as grep and in word processors such as Microsoft Word.

Templates are always defined in three parts: a prefix regex, a target regex, and a postfix regex. For prices, the target regex is as above, the prefix would look for strings such as "price:" and the postfix could be empty. The idea is that some clues about an attribute come from the attribute value itself and some come from the surrounding text.

If a regular expression for an attribute matches the text exactly once, then we can pull out the portion of the text that is the value of the attribute. If there aremiss-match, it means that was a default value or given attribute missing; but if there are several matches, we

Scanned with CamScanner

should a process to choose among them. One strategy is to have several templates for each attribute, ordered by priority.

For example, the top-priority example for price might look for the prefix "our price:" if that is not found, we look for the prefix "price:" and if that is not found, the empty prefix. Another strategy is to require all the matches and find some way to choose among them.

One step up from attribute-based extraction systems are **relational extraction** systems, which deal with multiple objects and also the relations among them. Thus, when these systems see the text "Rs.249.99," they need to determine not just that it is a price, but also which object has that price.

There are relational-based extraction system is FASTUS, which handles news stories about corporate mergers and acquisitions. It can read the story Bridgestone Sports Co. said Friday it has set up a joint venture in Taiwan with a local concern and a Japanese trading house to produce golf clubs to be shipped to Japan. And extract the relations:

$e \in$ Joint Ventures $\wedge$ Product(e, "*golf* clubs") $\wedge$ Date(e, "Friday")

$\wedge$ Member (e, "Bridgestone Sports Co") $\wedge$ Member (e, "a local concern")

$\wedge$ Member (e, "a Japanese trading house").

A relational extraction system can be built as a series of **cascaded finite-state transducers.**

So the system consists of a series of small, efficient finite-state automata (FSAs), where each automaton receives text as input, transfers the text into a different format, and passes it along with the next automaton. FASTUS consists of five stages:

1. Tokenization
2. Complex-word handling
3. Basic-group handling
4. Complex-phrase handling
5. Structure merging

FASTUS's first stage is **tokenization**, which segments the stream of characters into tokens (words, numbers, and punctuation). For English, tokenization can be simple; just separating characters at white space or punctuation does a fairly good job. Some tokenizes also deal with markup languages such as HTML, SGML, and XML.

The second stage handles **complex words**, including collocations such as "set up" and "joint venture," as well as proper names such as "Bridgestone Sports Co." These are recognized by a combination of lexical entries and finite-state grammar rules. For example, a company name might be recognized by the rule

Capitalized Word+ ("Company" | "Co" | "Inc" | "Ltd").

Scanned with CamScanner

The third stage handles **basic groups**, meaning noun groups and verb groups. The idea is to chunk these into units which will be managed by the later stages. But here we have simple rules that only approximate the complexity of English, but have the advantage of being representable by finite state automata. The example sentence would emerge from this stage as the following sequence of tagged groups:

1 NG: Bridgestone Sports Co. 2 VG: said 11 CJ: and

3 NG: Friday 12 NG: a Japanese trading house

4 NG: it 13 VG: to produce

5 VG: had set up 14 NG: golf clubs

6 NG: a joint venture 15 VG: to be shipped

7 PR: in 16 PR: to

8 NG: Taiwan 17 NG: Japan

9 PR: with

10 NG: a local concern

Here NG stands for noun group, VG stands for verb group, PR stands for preposition, and CJ stands for conjunction.

The fourth stage combines the basic groups into **complex phrases**. Again, the aim is to have rules that are finite-state and therefore it can be processed quickly, and will result in unambiguous (or nearly unambiguous) output phrases. One type of combination rule deals with domain-specific events. For example, the rule

Company+ SetUpJointVenture ("with" Company+)?

Captures one way to describe the formation of a joint venture. This stage is the first one inthe cascade where the output is placed into a database template as well as being placed in theoutput stream.

The final stage **merges structures** which were built up in the previous step. If the next sentence says "The joint venture will start production in January," then this step will notice that there are two references to a joint venture, and that they should be merged into one. This is an instance of the **identity uncertainty problem.**

In general, finite-state template-based information extraction works well for a restricted domain in which it is possible to predetermine, and how they will be mentioned. The cascaded transducer model helps modularize the necessary knowledge, easing construction of the system. These systems work especially well when they are reverse engineering text that has been generated by a program. For example, a shopping site on the Web is generated by a program that takes database entries and formats them into Webpages; a template-based extractor then recovers the original database.

Scanned with CamScanner

Finite-state information extraction is less successful at recovering information in highly variable format, such as text written by humans on a variety of subjects.

3.4.2 Probabilistic model for information extraction.

Probabilistic language models based on n-grams recover a surprising quantity of data about a language. They can perform well on such diverse tasks as language identification, spelling correction, genre classification, and named-entity recognition.

These language models have millions of features, therefore feature selection and preprocessing of the data to reduce noise is important.

When information extraction must be attempted from noisy or varied input, simple finite state approaches fare poorly. It is too hard to get all the rules and their priorities right; it is better to use a probabilistic model rather than a rule-based model. The simplest probabilistic model for sequences with hidden state is the **hidden Markov model, or HMM.**

To apply HMMs to information extraction, we can either build one big HMM for all the attributes or build a separate HMM for each attribute. We'll do the second. The observations are the words of the text, and the hidden states are whether we are within the target, prefix, or postfix part of the attribute template, or in the background (not part of a template).

For example, here is a brief text and the most probable (Viterbi) path for that text for two HMMs, one trained to acknowledge the speaker in a talk announcement, and one trained to acknowledge dates. The "-" indicates a background state:

Text: There will be a seminar by Dr. Andrew McCallum on Friday

Speaker: - - - - PRE PRE TARGET TARGETTARGET POST -

Date: - - - - - - - - - PRE TARGET

HMMs have two big advantages over FSAs for extraction.

1. HMMs are probabilistic, and thus tolerant to noise. In a regular expression, if a single expected character is missing, the regex fails to match; with HMMs there is graceful degradation with missing characters/words, and we get a probability indicating the degree of match, not just a Boolean match/fail.

Scanned with CamScanner

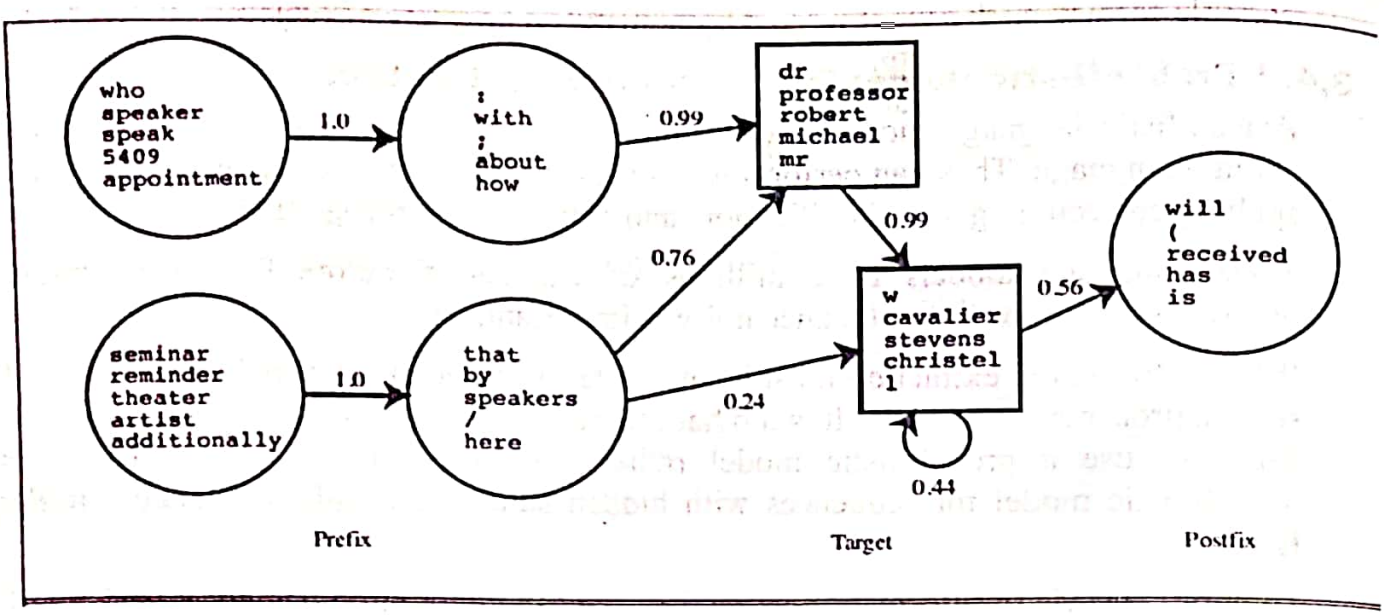


Figure: Hidden Markow model for the speaker of a talk announcement. The two aquare state are the target, the four circle to the left are the prefix, and the one on the right is the postfix. For each state, only a few of the high-probabilty words are shown. From Freitag and McCallum (2000).

2. HMMs can be trained from data; they don't require laborious engineering of templates, andthus they can more easily be kept up to date as text changes over time.

Note that we have assumed a certain level of structure in our HMM templates: they all consist of one or more target states, and any prefix states must precede the targets, postfix states most follow the targets, and other states must be background. This structure makesit easier to learn HMMs from examples.

For example, the word "Friday" would have high probability in one or more of the target states of the date HMM, and lower probability elsewhere.With sufficient training data, the HMM automatically learns a structure of dates that we find intuitive: the date HMM might have one target state in which the high-probability words are "Monday," "Tuesday," etc., and which has a high-probability transition to a target statewith words "Jan", "January," "Feb," etc.

Figure shows the HMM for the speaker of a vtalk announcement, as learned from data. The prefix covers expressions such as "Speaker:"and "seminar by," and the target has one state that covers titles and first names and anotherstate that covers initials and last names.

Once the HMMs have been learned, we can apply them to a text, using the Viterbialgorithm to find the most likely path through the HMM states.One approach is to applyeach attribute HMM separately; in this case you would expect most of the HMMs to

Scanned with CamScanner

spendmost of their time in background states. This is appropriate when the extraction is sparse when the number of extracted words is small compared to the length of the text.

The other approach is to combine all the individual attributes into one big HMM, whichwould then find a path that wanders through different target attributes, first finding a speakertarget, then a date target, etc. Separate HMMs are better when we expect just one of eachattribute in a text and one big HMM is better when the texts are more free-form and densewith attributes. With either approach, in the end we have a collection of target attributeobservations, and have to decide what to do with them. If every expected attribute has onetarget filler then the decision is easy: we have an instance of the desired relation. If thereare multiple fillers, we need to decide which to choose, as we discussed with template-basedsystems. HMMs have the advantage of supplying probability numbers that can help makethe choice. If some targets are missing, we need to decide if this is an instance of the desiredrelation at all, or if the targets found are false positives. A machine learning algorithm can betrained to make this choice.

3.5 Examples: Applications of Natural Language Processing:

A few examples of NLP that people use every day are:

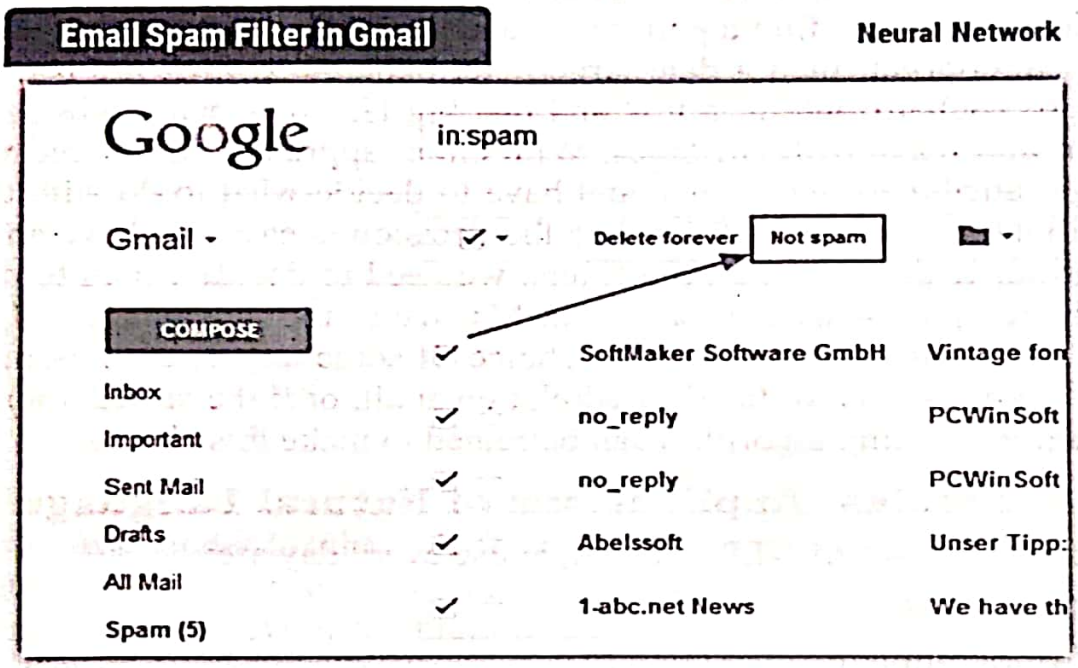
- Spell check
- Autocomplete
- Voice text messaging
- Search results
- Spam filters
- Related keywords on search engines
- Siri, Alexa, or Google Assistant

In any case, the computer is able to identify the appropriate word, phrase, or response by using context clues, the same way that any human would. Conceptually, it's a fairly straightforward technology.

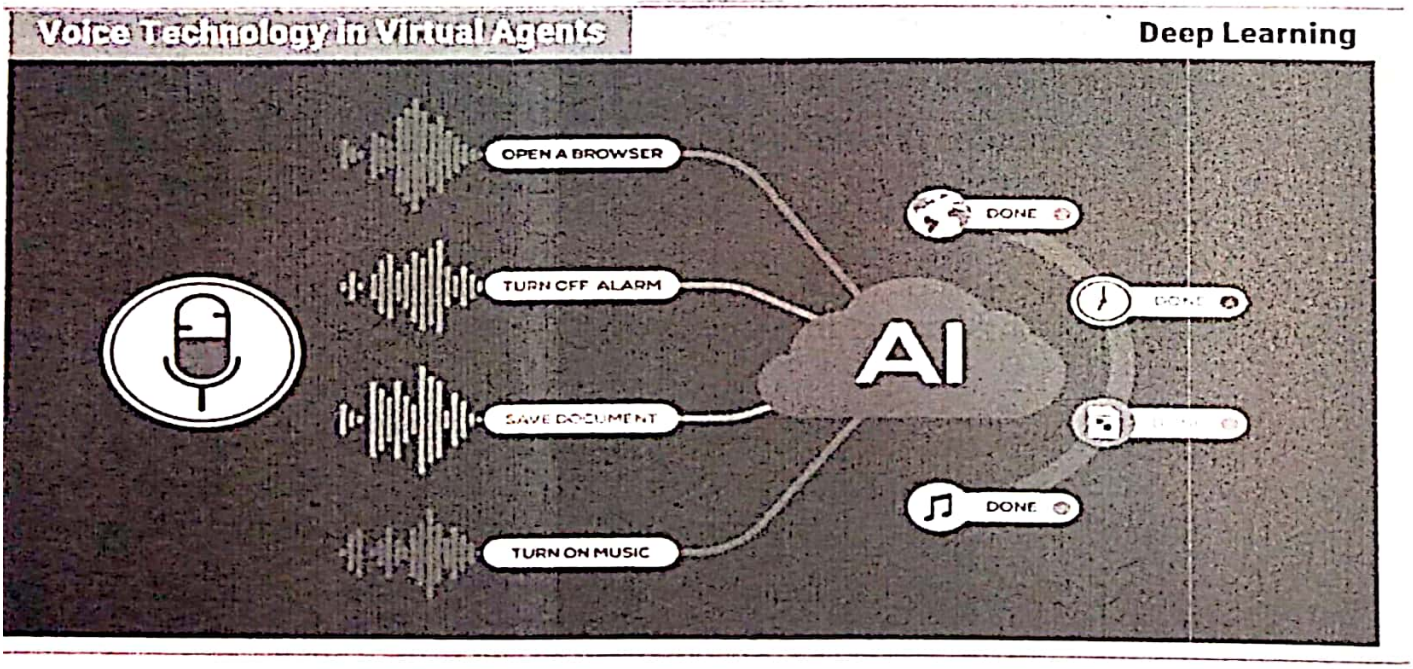
The best-known example of NLP, smart assistants such as Siri, Alexa and Cortana have become increasingly integrated into our lives

Scanned with CamScanner

Natural Language Processing

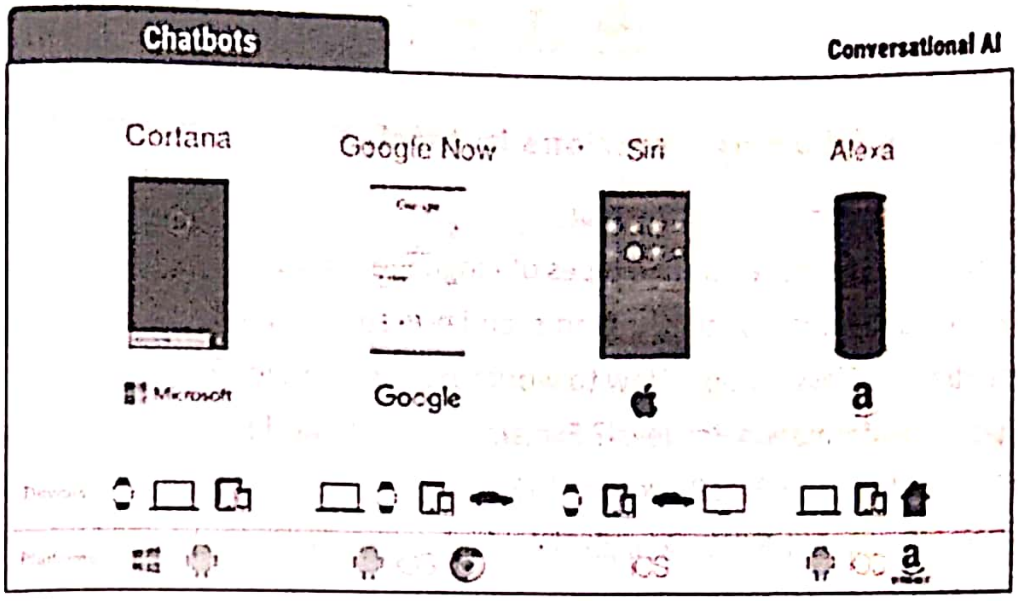


Case Study: Automated Voice Assistants, Chat bots.



Scanned with CamScanner

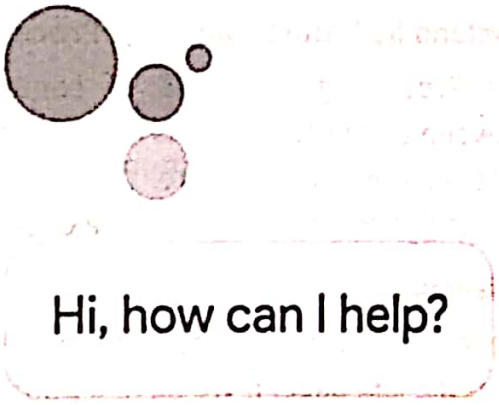
Chat bots.



Automate support:

Chatbots are nothing new, but advancements in NLP have increased their usefulness to the point that live agents no longer need to be the first point of communication for some customers. Some features of chatbots include being able to help users navigate support articles and knowledge bases, order products or services, and manage accounts.

- **Chatbots**: To provide a better customer support service, companies have started using chatbots for 24/7 service. Chatbots helps resolve the basic queries of customers. If a chatbot is not able to resolve any query, then it forwards it to the support team, while still engaging the customer. It helps make customers feel that the customer support team is quickly attending them. With the help of chatbots, companies have become capable of building cordial relations with customers. It is only possible with the help of Natural Language Processing.



Scanned with CamScanner

Exercises

❖ Answer the following Questions in brief.

1. Define natural language processing.
2. Define language model. Define types of language model.
3. What are unigrams, bigrams, trigrams, and n-grams in NLP?
4. Describe text Classification. How to work text compression?
5. What is an information Retrieval? Explain page rank and HITS Algorithms.
6. Difference between PageRank and HITS.
7. Define information Extraction. Explain finite state automata (FSA) for information extraction.
8. Define Probabilistic model.

❖ Short Questions with Answer:

1. What is N-gram character model?

Ans: An n-gram model is a technique of counting sequences of characters or words that allows us to support rich pattern discovery in text. ... In other words, it tries to capture patterns of sequences (characters or words next to each other) while being sensitive to contextual relations (characters or words near each other).

2. What is the corpus in NLP?

Ans: Corpus or corpora (plural), is a collection of the text of a similar type, for example, movie reviews, social media posts, etc.

3. What do you understand by Natural Language Processing?

Ans: Natural Language Processing is a field of computer science that deals with communication between computer systems and humans.
It is used to create automated software that helps understand human spoken languages to extract useful information from the data it gets in the form of audio.

4. What is an n-gram language model?

Ans: An N-gram language model predicts the probability of a given N-gram within any sequence of words in the language. If we have a good N-gram model, we can predict p(w | h) – what is the probability of seeing the word w given a history of previous words h – where the history contains n-1 words.

Scanned with CamScanner

N-gram refers to a number *n* of sequential items. A word 1-gram (uni-gram) for the sentence *"this is true"* would be [["this"],["is"],["true"]]; A 2-gram (bi-gram) would be: [["this", "is"], ["is", "true"]].

5. What is text classification in AI?

Ans: Text classification is the process of classifying documents into predefined categories based on their content. It is the automated assignment of natural language texts to predefined categories.

6. Define information retrieval.

Ans: Information retrieval (IR) is finding material (usually documents) of an unstructured nature (usually text) that satisfies an information need from within large collections (usually stored on computers).

7. What are the areas of AI for information retrieval?

Ans: Concepts surveyed include pattern recognition, representation, problem solving and planning, heuristics, and learning. The paper concludes with an outline of areas for further research on artificial intelligence in information retrieval systems.

8. Explain about HITS algorithms

Ans:

- Hypertext induced Topic selection is a link analysis method developed by John Kleinberg in 1999 using Hub and Authority scores.

  Two sets of inter-related pages:

- Hub Pages-good lists of links on a subject
- Authority pages-occur recurrently on good hubs for the subjects.

  The HITS algorithm

  $H(x) <- \sum a(y)$

  $A(x) <- \sum h(y)$

9. What is Recall?

Ans: Recall is the ratio of the number of relevant documents retrieved to the total number of relevant documents retrieved.

10. What is precision?

Ans: Precision is the ratio of the number of relevant documents retrieved to the total number of documents retrieved.

Scanned with CamScanner

11. Define authorities?

Ans: Authorities are pages that are recognized as providing significant, trustworthy and useful information on a topic. In-degree is one simple measure of authority. However Indegree treats all links as equal.

12. Define hubs.

Ans: Hubs are index pages that provide lots of useful links to relevant content pages. Hub pages for IR are included in the home page.

❖ Multipal choice Queations - MCQs:

1. What is full form of NLP?
   A. Nature Language Understanding
   B. Natural Long Processed
   C. Natural Language Processing
   D. None of the Above

2. Natural language processing can be divided into the two subfields of:
   A. context and expectations
   B. generation and understanding
   C. semantics of pragmatics
   D. recognition and synthesis

3. The area of AI that investigates methods of facilitating communication between people and computers is:
   A. natural language processing
   B. symbolic processing
   C. decision support
   D. robotics

4. 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

5. Given a sound clip of a person or people speaking, determine the textual representation of the speech.
   A. Text-to-speech
   B. Speech-to-text
   C. Both A and B
   D. None of the Above

6. What is Machine Translation?
   A. Converts one human language to another
   B. Converts human language to machine language
   C. Converts any human language to English
   D. Converts Machine language to human language

Scanned with CamScanner

7. The n-gram models includes _____.

A. Unigram
B. trigram
C. Bigram
D. All of the Above

8. Data Compression means to the file size.

A. Increase
B. Decrease
C. Can't say
D. none of the above

9. Data compression and encryption both work on binary

A. false
B. true

10. 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

11. Why data compressed?

A. To optimise the data
B. To reduce secondary storage space
C. To reduce packet congestion on network
D. Both (b) and (c)

12. What is a type of data compression?

A. Resolution
B. Zipping
C. Inputting
D. Caching

13. Data compression involves

A. compression only
B. Reconstruction only
C. Both compression and Reconstruction
D. None of the above

14. Compression is the method which eliminates the data which is not noticeable and compression does not eliminate the data which is not

A. Lossless, Lossy
B. Lossy, lossless
C. None of these
D. None of the above

15. IR objective is to retrieve _____ the relevant documents.

A. All
B. Specified number
C. half
D. None of the above

16. Full form of HITS is _____

A. Hyperlink Indicated Topic Search
B. Hyperlink Induced Topic Search

Scanned with CamScanner

C. Hyphen Indeed Topic Search D. Hyperlink Induced Transfer Search

17. _____ played important roles in developing our understanding of Web information retrieval.

A. Authority B. Hub

C. HITS D. Both PageRank and HITS

18. Bag of Words in text preprocessing is a _____

A. Feature scaling technique B. Feature extraction technique

C. Feature selection technique D. None

Answer:

1. Ans.: C Natural Language Processing
2. Ans.: B. generation and understanding
3. Ans.: A. natural language processing
4. Ans.: B. Speech and Written Text
5. Ans.: B
6. Ans.: A. converts one human language to another
7. Ans.: D. All of the Above
8. Ans.: B. Decrease
9. Ans.: B. true
10. Ans.: C. To reduce the size of data to save space
11. Ans.: D. Both (b) and (c)
12. Ans.: B. Zipping
13. Ans.: C. Both compression and Reconstruction
14. Ans.: B. Lossy, lossless
15. Ans.: A. All
16. Ans.: B. Hyperlink Induced Topic Search
17. Ans.: D. Both PageRank and HIT
18. Ans.: B. Feature extraction technique

Scanned with CamScanner

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.

Scanned with CamScanner

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

Scanned with CamScanner

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.

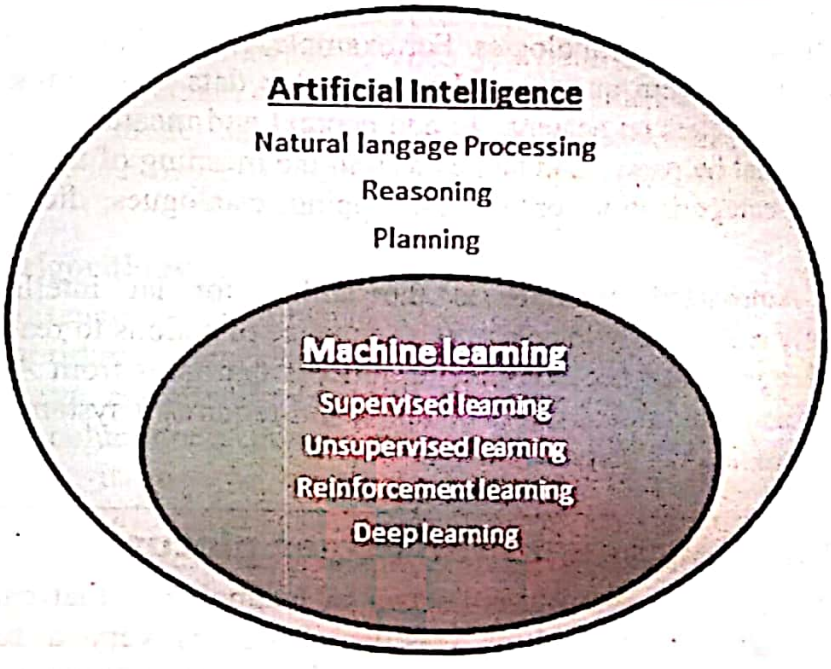


Fig: Relationship between AI and ML

Scanned with CamScanner

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,

Scanned with CamScanner

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.

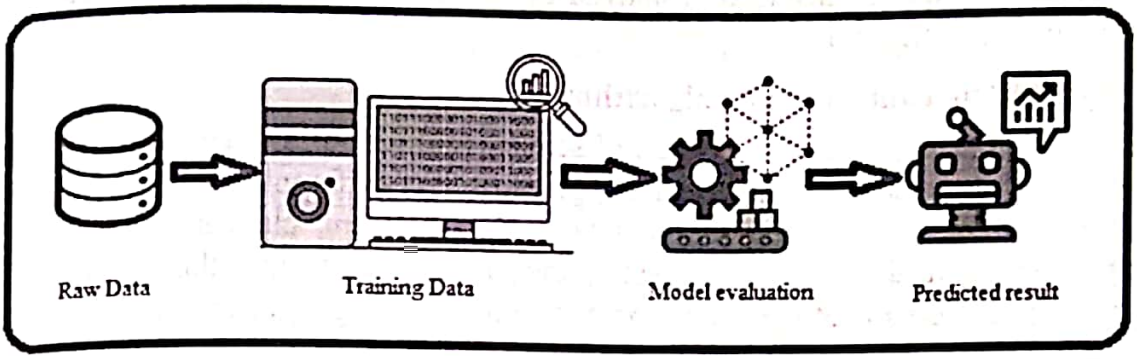


Fig: Workings of Machine learning

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

Scanned with CamScanner

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.

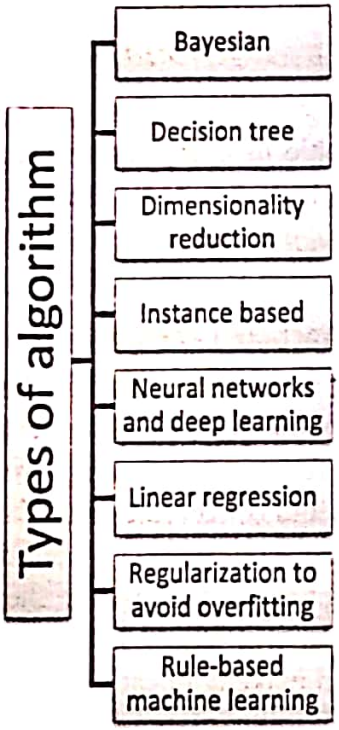


Fig: Types of Algorithms

➢ 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.

Scanned with CamScanner

➢ **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

Scanned with CamScanner

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.

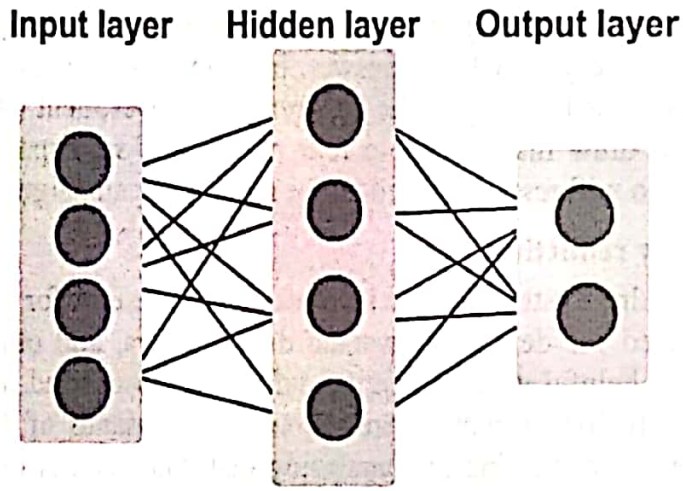


Figure: shows the formation of a neural network. Each layer of neural network filters and change details before passing it will be in the next layer.

➢ **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

Scanned with CamScanner

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

Scanned with CamScanner

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.

Scanned with CamScanner

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:

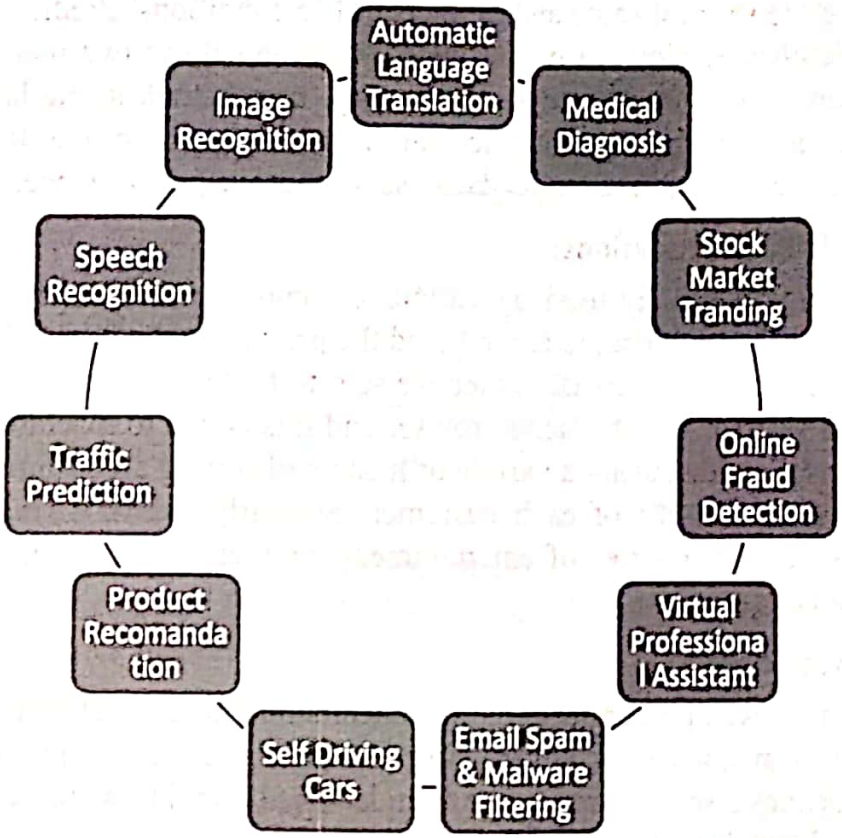


Fig: Application Areas of Machine learning

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

Scanned with CamScanner

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.

Scanned with CamScanner

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.

Scanned with CamScanner

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.

Scanned with CamScanner

➢ **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

Scanned with CamScanner

- ✓ 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:

Scanned with CamScanner

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:

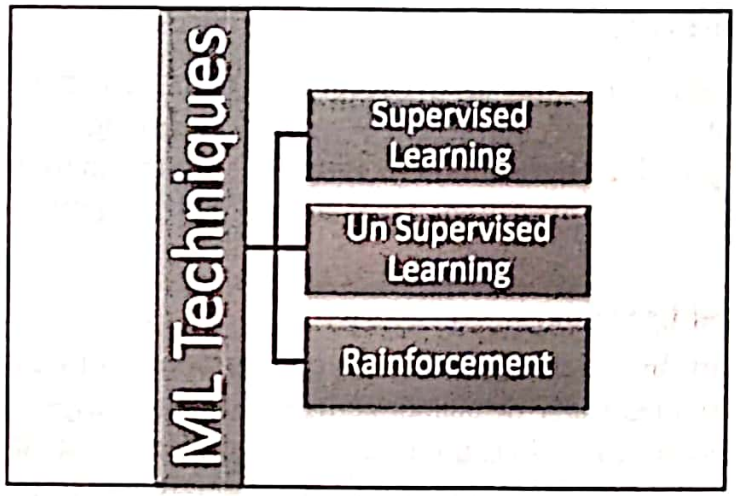


Fig: Types of Machine Learning Techniques

➢ **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

Scanned with CamScanner

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

Scanned with CamScanner

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.

Scanned with CamScanner

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.

Scanned with CamScanner

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. "

Scanned with CamScanner

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 S 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

Scanned with CamScanner

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

Scanned with CamScanner

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

Scanned with CamScanner

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

Scanned with CamScanner

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

Scanned with CamScanner

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

☺ ☺ ☺ ☺ ☺

Scanned with CamScanner