

**Title: Artificial Intelligence for Engineers**

**Authors Name:** Er. Karan Kumar

## ARTIFICIAL INTELLIGENCE FOR ENGINEERS

The students will be able to	Blooms Taxonomy
CO1 Understand the evolution and various approaches of AI	K2
CO2 Understand data storage, processing, visualization, and its use in regression, Clustering etc.	K2
CO3 Understand natural language processing and chatbots	K2
CO4 Understand the concepts of neural networks	K2
CO5 Understand the concepts of face, object, speech recognition and robots	K2

### Syllabus

**Unit 1 :** An overview to AI: The evolution of AI to the present, Various approaches to AI, What should all engineers know about AI?, Other emerging technologies, AI and ethical concerns

**Unit 2:** Data & Algorithms: History Of Data, Data Storage And Importance of Data and its Acquisition, The Stages of data processing, Data Visualization, Regression, Prediction & Classification, Clustering & Recommender Systems

**Unit 3:** Natural Language Processing: Speech recognition, Natural language understanding, Natural language generation, Chatbots, Machine Translation

**Unit 4:** Artificial Neural Networks: Deep Learning, Recurrent Neural Networks, Convolutional Neural Networks, The Universal Approximation Theorem, Generative Adversarial Networks

**Unit 5:** Applications: Image and face recognition, Object recognition, Speech Recognition besides Computer Vision, Robots, Applications

**Index:**

1. An overview to AI .....	04-22
1.1. The evolution of AI to the present.....	07
1.2. Various approaches to AI.....	09
1.3. What should all engineers know about AI?.....	10
1.4. Other merging technologies.....	13
1.5. AI and ethical concerns.....	20
2. Data & Algorithms.....	23-37
2.1. History Of Data.....	24
2.2. Data Storage And Importance of Data and its Acquisition.....	25
2.3. The Stages of data processing.....	27
2.4. Data Visualization.....	28
2.5. Regression, Prediction & Classification.....	30
2.6. Clustering & Recommender Systems.....	33
3. Natural Language Processing.....	37-49
3.1. Speech recognition.....	39
3.2. Natural language understanding .....	42
3.3. Natural language generation.....	45
3.4. Chatbots.....	45
3.5. Machine Translation.....	47
4. Artificial Neural Networks.....	50-62
4.1. Deep Learning.....	52
4.2. Recurrent Neural Networks.....	54
4.3. Convolutional Neural Networks.....	57
4.4. The Universal Approximation Theorem .....	58
4.5. Generative Adversarial Networks.....	59
5. Applications.....	63-75
5.1. Image and face recognition.....	63
5.2. Object recognition.....	66
5.3. Speech Recognition besides Computer Vision.....	67
5.4. Robots.....	69
5.5. Applications.....	72

## Unit-1

### An overview to AI

"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."

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.

### 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.

### History of AI

Here is the history of AI during 20th century

Year    Milestone / Innovation

- 1923    Karel Čapek play named "Rossum's Universal Robots" (RUR) opens in London, first use of the word "robot" in English.
- 1943    Foundations for neural networks laid.
- 1945    Isaac Asimov, a Columbia University alumni, coined the term Robotics.

- 1950 Alan Turing introduced Turing Test for evaluation of intelligence and published Computing Machinery and Intelligence. Claude Shannon published Detailed Analysis of Chess Playing as a search.
- 1956 John McCarthy coined the term Artificial Intelligence. Demonstration of the first running AI program at Carnegie Mellon University.
- 1958 John McCarthy invents LISP programming language for AI.
- 1964 Danny Bobrow's dissertation at MIT showed that computers can understand natural language well enough to solve algebra word problems correctly.
- 1965 Joseph Weizenbaum at MIT built ELIZA, an interactive problem that carries on a dialogue in English.
- 1969 Scientists at Stanford Research Institute Developed Shakey, a robot, equipped with locomotion, perception, and problem solving.
- 1973 The Assembly Robotics group at Edinburgh University built Freddy, the Famous Scottish Robot, capable of using vision to locate and assemble models.
- 1979 The first computer-controlled autonomous vehicle, Stanford Cart, was built.
- 1985 Harold Cohen created and demonstrated the drawing program, Aaron.
- 1990 Major advances in all areas of AI –
  - Significant demonstrations in machine learning
  - Case-based reasoning
  - Multi-agent planning
  - Scheduling
  - Data mining, Web Crawler
  - natural language understanding and translation
  - Vision, Virtual Reality
  - Games
- 1997 The Deep Blue Chess Program beats the then world chess champion, Garry Kasparov.

2000 Interactive robot pets become commercially available. MIT displays Kismet, a robot with a face that expresses emotions. The robot Nomad explores remote regions of Antarctica and locates meteorites.

## Applications of AI

AI has been dominant in various fields such as –

- **Gaming** – AI plays crucial role in strategic games such as chess, poker, tic-tac-toe, etc., where machine can think of large number of possible positions based on heuristic knowledge.
- **Natural Language Processing** – It is possible to interact with the computer that understands natural language spoken by humans.
- **Expert Systems** – There are some applications which integrate machine, software, and special information to impart reasoning and advising. They provide explanation and advice to the users.
- **Vision Systems** – These systems understand, interpret, and comprehend visual input on the computer. For example,
  - A spying aeroplane takes photographs, which are used to figure out spatial information or map of the areas.
  - Doctors use clinical expert system to diagnose the patient.
  - Police use computer software that can recognize the face of criminal with the stored portrait made by forensic artist.
- **Speech Recognition** – Some intelligent systems are capable of hearing and comprehending the language in terms of sentences and their meanings while a human talks to it. It can handle different accents, slang words, noise in the background, change in human's noise due to cold, etc.
- **Handwriting Recognition** – The handwriting recognition software reads the text written on paper by a pen or on screen by a stylus. It can recognize the shapes of the letters and convert it into editable text.
- **Intelligent Robots** – Robots are able to perform the tasks given by a human. They have sensors to detect physical data from the real world such as light, heat, temperature, movement, sound, bump, and pressure. They have efficient processors, multiple sensors and huge memory, to exhibit intelligence. In addition, they are capable of learning from their mistakes and they can adapt to the new environment.

## 1.1. The evolution of AI to the present

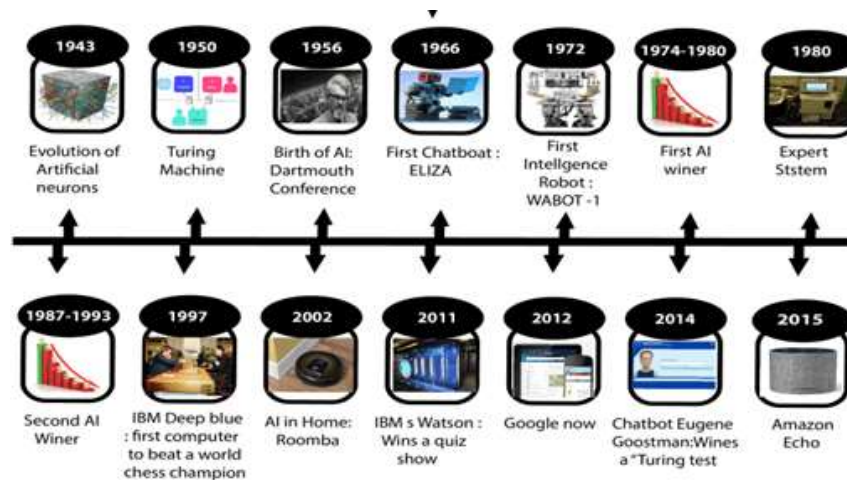


Figure 1.1: Evolution of AI

### Maturation of Artificial Intelligence (1943-1952)

- **Year 1943:** The first work which is now recognized as AI was done by Warren McCulloch and Walter Pitts 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:** 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 found new and more elegant proofs for some theorems.
- **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)

- **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.

## 1.2. Various approaches to AI

Four approaches have been followed. As one might expect, a tension exists between approaches centred on humans and approaches centered around rationality. A Human centred 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

### Acting humanly: The Turing Test approach

**The Turing Test**, proposed by Alan Turing (1950), was designed to provide a satisfactory operational definition of intelligence. Turing defined intelligent behavior as the ability to achieve human-level performance in all cognitive tasks, sufficient to fool an interrogator. Roughly speaking, the test he proposed is that the computer should be interrogated by a human via a teletype, and passes the test if the interrogator cannot tell if there is a computer or a human at the

other end. Chapter 26 discusses the details of the test, and whether or not a computer is really intelligent if it passes. For now, programming a computer to pass the test provides plenty to work on. The computer would need to possess the following capabilities:

- natural language processing to enable it to communicate successfully in English (or some other human language);
- knowledge representation to store information provided before or during the interrogation;
- automated reasoning to use the stored information to answer questions and to draw new conclusions;
- machine learning to adapt to new circumstances and to detect and extrapolate patterns

**Thinking humanly:** If we are going to say that a given program thinks like a human, we must have some way of determining how humans think. We need to get inside the actual workings of human minds.

**Thinking rationally:** The Greek philosopher Aristotle was one of the first to attempt to codify "right thinking," that is, irrefutable reasoning processes. His famous **syllogisms** provided patterns for argument structures that always gave correct conclusions given correct premises.

**Acting rationally:** Acting rationally means acting so as to achieve one's goals, given one's beliefs. An **agent** is just something that perceives and acts. (This may be an unusual use of the word, but you will get used to it.) In this approach, AI is viewed as the study and construction of rational agents.

### 1.3. What should all engineers know about AI?

An AI engineer builds AI models using machine learning algorithms and deep learning neural networks to draw business insights, which can be used to make business decisions that affect the entire organization. These engineers also create weak or strong AIs, depending on what goals they want to achieve.

AI engineers have a sound understanding of programming, software engineering, and data science. They use different tools and techniques so they can process data, as well as develop and maintain AI systems.

The next section of How to become an AI Engineer focuses on the responsibilities of an AI engineer.

**Responsibilities of an AI Engineer**

As an AI engineer, you need to perform certain tasks, such as develop, test, and deploy AI models through programming algorithms like random forest, logistic regression, linear regression, and so on.

**Responsibilities include:**

- Convert the machine learning models into application program interfaces (APIs) so that other applications can use it
- Build AI models from scratch and help the different components of the organization (such as product managers and stakeholders) understand what results they gain from the model
- Build data ingestion and data transformation infrastructure
- Automate infrastructure that the data science team uses
- Perform statistical analysis and tune the results so that the organization can make better-informed decisions
- Set up and manage AI development and product infrastructure
- Be a good team player, as coordinating with others is a must

**Skills Required to Become an AI Engineer**

Professionals who are finding how to become an AI engineer should also know about the skills required in this field. Some of them include:

- Programming Skills

The first skill required to become an AI engineer is programming. To become well-versed in AI, it's crucial to learn programming languages, such as Python, R, Java, and C++ to build and implement models.

- Linear Algebra, Probability, and Statistics

To understand and implement different AI models—such as Hidden Markov models, Naive Bayes, Gaussian mixture models, and linear discriminant analysis—you must have detailed knowledge of linear algebra, probability, and statistics.

- Spark and Big Data Technologies

AI engineers work with large volumes of data, which could be streaming or real-time production-level data in terabytes or petabytes. For such data, these engineers need to know about Spark and other big data technologies to make sense of it. Along with Apache Spark, one can also use other big data technologies, such as Hadoop, Cassandra, and MongoDB.

- Algorithms and Frameworks

Understanding how machine learning algorithms like linear regression, KNN, Naive Bayes, Support Vector Machine, and others work will help you implement machine learning models with ease. Additionally, to build AI models with unstructured data, you should understand deep learning algorithms (like a convolutional neural network, recurrent neural network, and generative adversarial network) and implement them using a framework. Some of the frameworks used in artificial intelligence are PyTorch, Theano, TensorFlow, and Caffe.

- Communication and Problem-solving Skills

AI engineers need to communicate correctly to pitch their products and ideas to stakeholders. They should also have excellent problem-solving skills to resolve obstacles for decision making and drawing helpful business insights.

Let us explore the career and roles in AI in the next section of the How to become an AI Engineer article.

- AI Engineer Salary

According to Glassdoor, the average annual salary of an AI engineer is \$114,121 in the United States and ₹765,353 in India. The salary may differ in several organizations and with the knowledge and expertise, you bring to the table.

- Career in AI

Since several industries around the world use AI to some degree or the other, including healthcare and education, there has been exponential growth in the career opportunities within the field of AI. Some of these job roles are:

- AI Developer

An AI developer works closely with electrical engineers and develops software to create artificially intelligent robots.

- AI Architect

AI architects work closely with clients to provide constructive business and system integration services. They also create and maintain the entire architecture.

- Machine Learning Engineer

Machine learning engineers build predictive models using vast volumes of data. They have in-depth knowledge of machine learning algorithms, deep learning algorithms, and deep learning frameworks.

- Data Scientists

Data scientists collect, clean, analyze, and interpret large and complex datasets by leveraging both machine learning and predictive analytics.

- Business Intelligence Developer

They're responsible for designing, modeling, and analyzing complex data to identify the business and market trends.

#### 1.4. Other emerging technologies

- **RPA (Robotic Process Automation) RPA Training** Robotic process automation (or RPA) is a form of business process automation technology based on metaphorical software robots (bots) or on artificial intelligence (AI)/digital workers. It is sometimes referred to as software robotics.
- Automating repetitive tasks saves time and money. Robotic process automation bots expand the value of an automation platform by completing tasks faster, allowing employees to perform higher-value work.



Figure 1.2: Robotic Process Automation

- **Big Data**

Big Data is a collection of data that is huge in volume, yet growing exponentially with time. It is a data with so large size and complexity that none of traditional data management tools.

Examples: Social Media

Hadoop and Spark are the two most famous frameworks for solving Big Data problems.

#### Types of Big Data

**Structured:** Any data that can be stored, accessed and processed in the form of fixed format is termed as a 'structured' data.

**Examples :** An 'Employee' table in a database is an example of Structured Data

Employee_ID	Employee_Name	Gender	Department	Salary_In_lacs
2365	Rajesh Kulkarni	Male	Finance	650000
3398	Pratibha Joshi	Female	Admin	650000
7465	Shushil Roy	Male	Admin	500000
7500	Shubhojit Das	Male	Finance	500000
7699	Priya Sane	Female	Finance	550000

Table 1.1: Structured data

**Unstructured :** Any data with unknown form or the structure is classified as unstructured data. In addition to the size being huge, un-structured data poses multiple challenges in terms of its processing for deriving value out of it. A typical example of unstructured data is a heterogeneous data source containing a combination of simple text files, images, videos etc.

**Example:** Google search

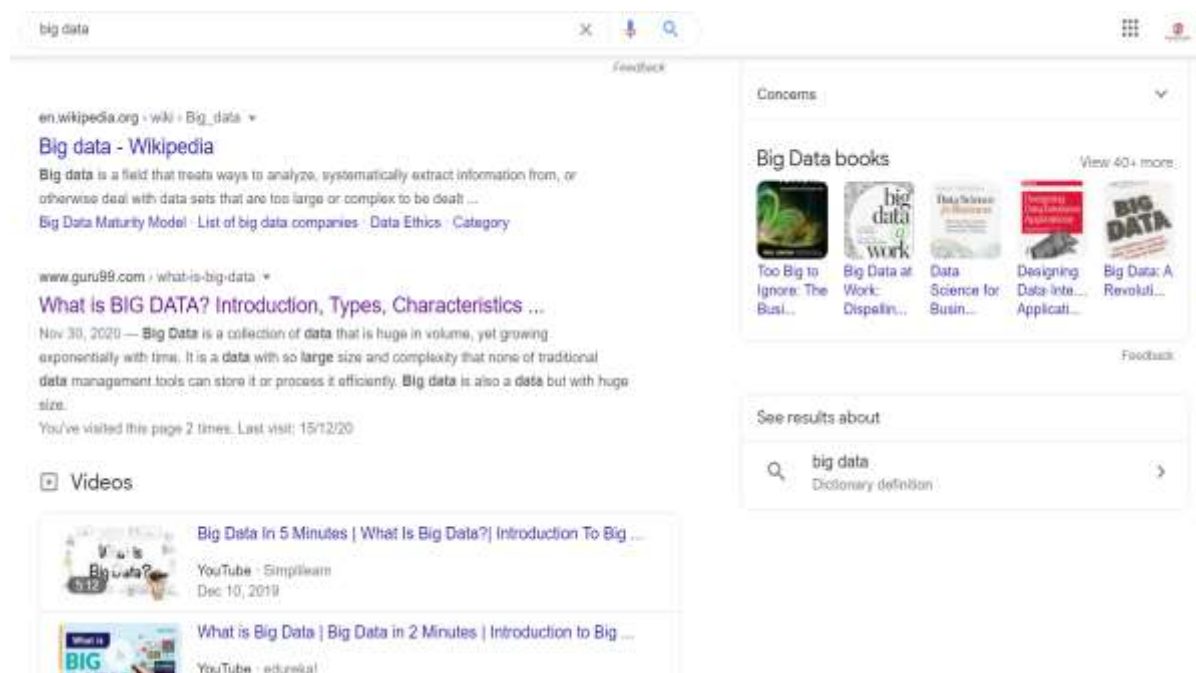


Figure 1.3: Unstructured data

**Semi-structured:** Semi-structured data can contain both the forms of data. We can see semi-structured data as a structured in form but it is actually not defined with e.g. a table definition in relational DBMS. Example of semi-structured data is a data represented in an XML file.

Examples of semi-structured: CSV but XML and JSON documents are semi structured documents, NoSQL databases are considered as semi structured.

### Characteristics Of Big Data

- (i) **Volume** – The name Big Data itself is related to a size which is enormous. Size of data plays a very crucial role in determining value out of data.
- (ii) **Variety** – Variety refers to heterogeneous sources and the nature of data, both structured and unstructured.
- (iii) **Velocity** – The term '**velocity**' refers to the speed of generation of data.
- (iv) **Variability** – This refers to the inconsistency which can be shown by the data at times, thus hampering the process of being able to handle and manage the data effectively.

### ➤ Intelligent Apps (I – Apps)

- I-Apps are pieces of software written for mobile devices based on artificial intelligence and machine learning technology, aimed at making everyday tasks easier.
- This involves tasks like organizing and prioritizing emails, scheduling meetings, logging interactions, content, etc. Some familiar examples of I-Apps are Chatbots and virtual assistants.

As these applications become more popular, they will come with the promise of jobs and fat paychecks.



Figure 1.4: Intelligent apps

➤ **Internet of Things (IoT)**

- The Internet of Things (IoT) describes the network of physical objects—“things”—that are embedded with sensors, software, and other technologies for the purpose of connecting and exchanging data with other devices and systems over the internet.
- These devices range from ordinary household objects to sophisticated industrial tools. With more than 7 billion connected IoT devices today, experts are expecting this number to grow to 10 billion by 2020 and 22 billion by 2025.

This includes everything from your:

- mobile phones,
- refrigerator,
- washing machines to almost everything that you can think of.

With IoT, we can have smart cities with optimized:

- traffic system,
- efficient waste management and
- energy use

So, start thinking of some new excuse for coming late to the office other than traffic.



Figure 1.5: The Internet of Things



### ➤ DevOps

**DevOps Training** This is the odd one out in the list. It is not a technology, but a methodology.

DevOps is a methodology that ensures that both the development and operations go hand in hand. DevOps cycle is picturized as an infinite loop representing the integration of developers and operation teams by:

- automating infrastructure,
- workflows and
- continuously measuring application performance.



Figure 1.6: DevOps

It is basically the process of continual improvement, so why not start with yourself.

### ➤ Angular and React

- Angular and React are JavaScript based Frameworks for creating modern web applications.
- Using React and Angular one can create a highly modular web app. So, you don't need to go through a lot of changes in your code base for adding a new feature.
- Angular and React also allows you to create a native mobile application with the same JS, CSS & HTML knowledge.
- Best part – Open source library with highly active community support.



Figure 1.7: Angular and React

### ➤ Cloud Computing

- cloud computing is the delivery of computing services—including servers, storage, databases, networking, software, analytics, and intelligence—over the Internet (“the cloud”) to offer faster innovation, flexible resources, and economies of scale.
- You typically pay only for cloud services you use, helping lower your operating costs, run your infrastructure more efficiently and scale as your business needs change

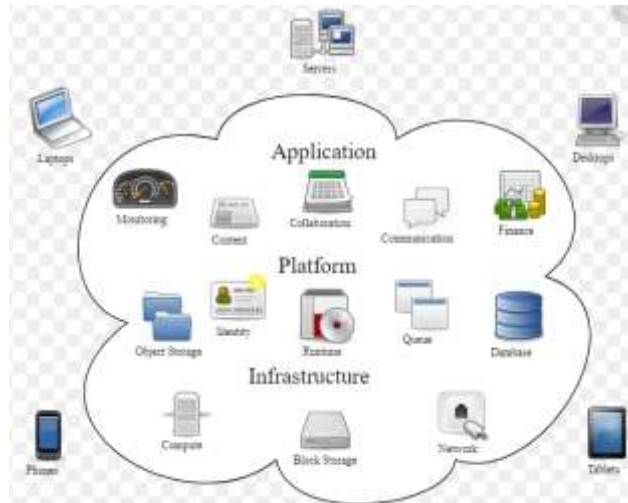


Figure 1.8: Cloud computing

### Types of cloud computing

**Public cloud:** Public clouds are owned and operated by a third-party cloud service providers, which deliver their computing resources like servers and storage over the Internet. Microsoft Azure is an example of a public cloud. With a public cloud, all hardware, software and other supporting infrastructure is owned and managed by the cloud provider.

**Private cloud:** A private cloud refers to cloud computing resources used exclusively by a single business or organisation. A private cloud can be physically located on the company’s on-site datacenter. Some companies also pay third-party service providers to host their private cloud. Example: HP Data Centers, Microsoft, Elastra-private cloud, and Ubuntu.

**Hybrid cloud:** Hybrid clouds combine public and private clouds, bound together by technology that allows data and applications to be shared between them. By allowing data and applications to move between private and public clouds, a hybrid cloud gives your business greater flexibility, more deployment options and helps optimise your existing infrastructure, security and compliance. Example : Amazon Web Services (AWS) or Microsoft Azure.

**Types of cloud services:**

- **IaaS (Infrastructure as Service):** This is the most common service model of cloud computing as it offers the fundamental infrastructure of virtual servers, network, operating systems and data storage drives. It allows for the flexibility, reliability and scalability that many businesses seek with the cloud, and removes the need for hardware in the office.
- **PaaS (Platform-as-a-Service):** This is where cloud computing providers deploy the infrastructure and software framework, but businesses can develop and run their own applications. Web applications can be created quickly and easily via PaaS, and the service is flexible and robust enough to support them.
- **SaaS (Software as a Service):** This cloud computing solution involves the deployment of software over the internet to various businesses who pay via subscription or a pay-per-use model. It is a valuable tool for CRM and for applications that need a lot of web or mobile access – such as mobile sales management software

**➤ Augmented Reality and Virtual Reality**

- Virtual is real! VR and AR, the twin technologies that let you experience things in virtual, that are extremely close to real, are today being used by businesses of all sizes and shapes. But the underlying technology can be quite complex.
- Medical students use AR technology to practice surgery in a controlled environment.
- VR on the other hand, opens up newer avenues for gaming and interactive marketing.

Whatever your interest might be, AR and VR are must-have skills if you want to ride the virtual wave!



Figure 1.9: Augmented Reality and Virtual Reality

### ➤ **Blockchain:**

Blockchain is a system of recording information in a way that makes it difficult or impossible to change, hack, or cheat the system. A blockchain is essentially a digital ledger of transactions that is duplicated and distributed across the entire network of computer systems on the blockchain. Blockchain, sometimes referred to as Distributed Ledger Technology (DLT), makes the history of any digital asset unalterable and transparent through the use of decentralization and cryptographic hashing.

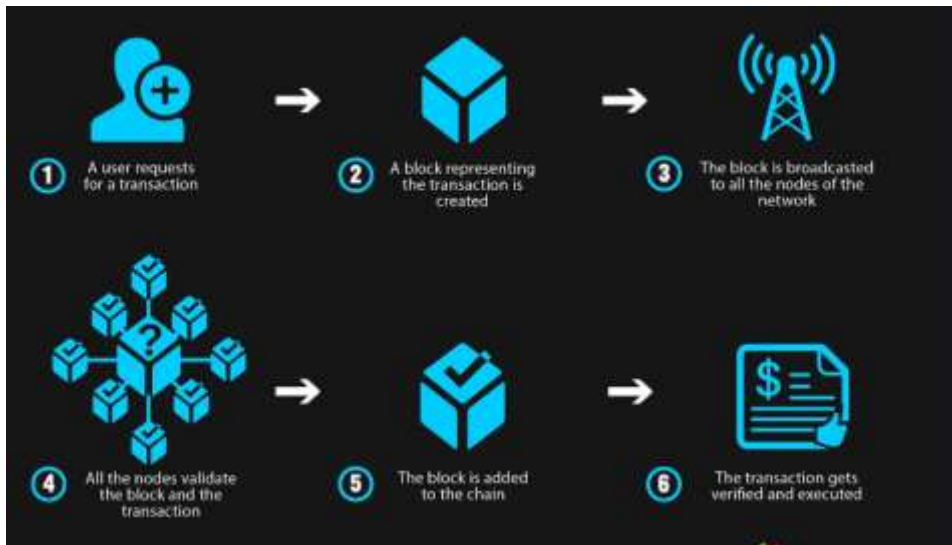


Figure 1.10: Working of blockchain

## 1.5. AI and ethical Concerns

### ➤ **Job Loss and Wealth Inequality**

One of the primary concerns people have with AI is future loss of jobs. Should we strive to fully develop and integrate AI into society if it means many people will lose their jobs — and quite possibly their livelihood?

### ➤ **AI is Imperfect — What if it Makes a Mistake?**

AIs are not immune to making mistakes and machine learning takes time to become useful. If trained well, using good data, then AIs can perform well. However, if we feed AIs bad data or make errors with internal programming, the AIs can be harmful. Teka Microsoft's AI chatbot, Tay, which was released on Twitter in 2016. In less than one day, due to the information it was receiving and learning from other Twitter users, the robot learned to spew racist slurs and Nazi propaganda.

Microsoft shut the chatbot down immediately since allowing it to live would have obviously damaged the company's reputation.

➤ **Should AI Systems Be Allowed to Kill?**

In this TEDx speech, Jay Tuck describes AIs as software that writes its own updates and renews itself. This means that, as programmed, the machine is not created to do what we want it to do — it does what it learns to do. Jay goes on to describe an incident with a robot called Tallon. Its computerized gun was jammed and open fired uncontrollably after an explosion killing 9 people and wounding 14 more.

➤ **Rogue AIs**

If there is a chance that intelligent machines can make mistakes, then it is within the realm of possibility that an AI can go rogue, or create unintended consequences from its actions in pursuing seemingly harmless goals. One scenario of an AI going rogue is what we've already seen in movies like The Terminator and TV shows where a super-intelligent centralized AI computer becomes self-aware and decides it doesn't want human control anymore.

➤ **Singularity and Keeping Control Over AIs**

Will AIs evolve to surpass human beings? What if they become smarter than humans and then try to control us? Will computers make humans obsolete? The point at which technology growth surpasses human intelligence is referred to as "technological singularity." Some believe this will signal the end of the human era and that it could occur as early as 2030 based on the pace of technological innovation. AIs leading to human extinction — it's easy to understand why the advancement of AI is scary to many people.

➤ **How Should We Treat AIs?**

Should robots be granted human rights or citizenship? If we evolve robots to the point that they are capable of "feeling," does that entitle them to rights similar to humans or animals? If robots are granted rights, then how do we rank their social status? This is one of the primary issues in "roboethics," a topic that was first raised by Isaac Asimov in 1942. In 2017, the Hanson Robotics humanoid robot, Sophia, was granted citizenship in Saudi Arabia. While some consider this to be more of a PR stunt than actual legal recognition, it does set an example of the type of rights AIs may be granted in the future.

➤ **AI Bias**

AI has become increasingly inherent in facial and voice recognition systems, some of which have real business implications and directly impact people. These systems are vulnerable to biases and

errors introduced by its human makers. Also, the data used to train these AI systems itself can have biases. For instance, facial recognition algorithms made by Microsoft, IBM and Megvii all had biases when detecting people's gender.

**Questions:**

1. What is Intelligence?
2. Describe the four categories under which AI is classified with examples.
3. Define Artificial Intelligence.
4. List the fields that form the basis for AI.
5. What are various approaches to AI.
6. What is emerging technologies? Give some examples.
7. What is the importance of ethical issue in AI?
8. Write the history of AI.
9. What are applications of AI?
10. What should all engineers know about AI?

## Unit-II

### Data & Algorithms

#### Data

In general, data is any set of characters that is gathered and translated for some purpose, usually analysis. If data is not put into context, it doesn't do anything to a human or computer.

There are multiple types of data. Some of the more common types of data include the following:

- Single character
- Boolean (true or false)
- Text (string)
- Number (integer or floating-point)
- Picture
- Sound
- Video

In a computer's storage, data is a series of bits (binary digits) that have the value one or zero. Data is processed by the CPU, which uses logical operations to produce new data (output) from source data (input).

#### Algorithms

Big data is data so large that it does not fit in the main memory of a single machine, and the need to process big data by efficient algorithms arises in Internet search, network traffic monitoring, machine learning, scientific computing, signal processing, and several other areas. This course will cover mathematically rigorous models for developing such algorithms, as well as some provable limitations of algorithms operating in those models. Some topics we will cover include:

- **Sketching and Streaming.** Extremely small-space data structures that can be updated on the fly in a fast-moving stream of input.
- **Dimensionality reduction.** General techniques and impossibility results for reducing data dimension while still preserving geometric structure.
- **Numerical linear algebra.** Algorithms for big matrices (e.g. a user/product rating matrix for Netflix or Amazon). Regression, low rank approximation, matrix completion, ...
- **Compressed sensing.** Recovery of (approximately) sparse signals based on few linear measurements.



- **External memory and cache-obliviousness.** Algorithms and data structures minimizing I/Os for data not fitting on memory but fitting on disk. B-trees, buffer trees, multiway mergesort, ...

## 2.1. History of Data

The history of big data starts many years before the present buzz around Big Data. Seventy years ago the first attempt to quantify the growth rate of data in the terms of volume of data was encountered. That has popularly been known as “information explosion“. We will be covering some major milestones in the evolution of “big data”.

1944: Fremont Rider, based upon his observation, speculated that Yale Library in 2040 will have “approximately 200,000,000 volumes, which will occupy over 6,000 miles of shelf. From 1944 to 1980, many articles and presentations were presented that observed the ‘information explosion’ and the arising needs for storage capacity.

1980: In 1980, the sociologist Charles Tilly uses the term big data in one sentence “none of the big questions has actually yielded to the bludgeoning of the big-data people.” in his article “The old-new social history and the new old social history”. But the term used in this sentence is not in the context of the present meaning of Big Data today.

1997: In 1977, Michael Cox and David Ellsworth published the article “Application-controlled demand paging for out-of-core visualization” in the Proceedings of the IEEE 8th conference on Visualization. The article uses the big data term in the sentence “Visualization provides an interesting challenge for computer systems: data sets are generally quite large, taxing the capacities of main memory, local disk, and even remote disk. We call this the problem of big data. When data sets do not fit in main memory (in core), or when they do not fit even on local disk, the most common solution is to acquire more resources.”.

1998: In 1998, John Mashey, who was Chief Scientist at SGI presented a paper titled “Big Data... and the Next Wave of Infrastrass.” at a USENIX meeting. John Mashey used this term in his various speeches and that’s why he got the credit for coining the term Big Data.



2000: In 2000, Francis Diebold presented a paper titled “‘Big Data’ Dynamic Factor Models for Macroeconomic Measurement and Forecasting” to the Eighth World Congress of the Econometric Society.

2001: In 2001, Doug Laney, who was an analyst with the Meta Group (Gartner), presented a research paper titled “3D Data Management: Controlling Data Volume, Velocity, and Variety.” The 3V’s have become the most accepted dimensions for defining big data.

2005: In 2005, Tim O’Reilly published his groundbreaking article “What is Web 2.0?”. In this article, Tim O’Reilly states that the “data is the next Intel inside”. O’Reilly Media explicitly used the term ‘Big Data’ to refer to the large sets of data which is almost impossible to handle and process using the traditional business intelligence tools.

In 2005 Yahoo used Hadoop to process petabytes of data which is now made open-source by Apache Software Foundation. Many companies are now using Hadoop to crunch Big Data.

So we can say that 2005 is the year that the Big data revolution has truly begun and the rest they say is history.

## 2.2. Data Storage And Importance of Data and its Acquisition

The systems, used for data acquisition are known as **data acquisition systems**. These data acquisition systems will perform the tasks such as conversion of data, storage of data, transmission of data and processing of data.

Data acquisition systems consider the following **analog signals**.

- Analog signals, which are obtained from the direct measurement of electrical quantities such as DC & AC voltages, DC & AC currents, resistance and etc.
- Analog signals, which are obtained from transducers such as LVDT, Thermocouple & etc.

### Types of Data Acquisition Systems

Data acquisition systems can be classified into the following **two types**.

- Analog Data Acquisition Systems
- Digital Data Acquisition Systems

Now, let us discuss about these two types of data acquisition systems one by one.

### ➤ Analog Data Acquisition Systems

The data acquisition systems, which can be operated with analog signals are known as **analog data acquisition systems**. Following are the blocks of analog data acquisition systems.

- **Transducer** – It converts physical quantities into electrical signals.
- **Signal conditioner** – It performs the functions like amplification and selection of desired portion of the signal.
- **Display device** – It displays the input signals for monitoring purpose.
- **Graphic recording instruments** – These can be used to make the record of input data permanently.
- **Magnetic tape instrumentation** – It is used for acquiring, storing & reproducing of input data.

### ➤ Digital Data Acquisition Systems

The data acquisition systems, which can be operated with digital signals are known as **digital data acquisition systems**. So, they use digital components for storing or displaying the information.

Mainly, the following **operations** take place in digital data acquisition.

- Acquisition of analog signals
- Conversion of analog signals into digital signals or digital data
- Processing of digital signals or digital data

Following are the blocks of **Digital data acquisition systems**.

- **Transducer** – It converts physical quantities into electrical signals.
- **Signal conditioner** – It performs the functions like amplification and selection of desired portion of the signal.
- **Multiplexer** – connects one of the multiple inputs to output. So, it acts as parallel to serial converter.
- **Analog to Digital Converter** – It converts the analog input into its equivalent digital output.
- **Display device** – It displays the data in digital format.

- **Digital Recorder** – It is used to record the data in digital format.

Data acquisition systems are being used in various applications such as biomedical and aerospace. So, we can choose either analog data acquisition systems or digital data acquisition systems based on the requirement.

### 2.3. The Stages of data Processing

Data processing occurs when data is collected and translated into usable information. Usually performed by a data scientist or team of data scientists, it is important for data processing to be done correctly as not to negatively affect the end product, or data output.

Data processing starts with data in its raw form and converts it into a more readable format (graphs, documents, etc.), giving it the form and context necessary to be interpreted by computers and utilized by employees throughout an organization.

Six stages of data processing

#### ➤ Data collection

Collecting data is the first step in data processing. Data is pulled from available sources, including data lakes and data warehouses. It is important that the data sources available are trustworthy and well-built so the data collected (and later used as information) is of the highest possible quality.

#### ➤ Data preparation

Once the data is collected, it then enters the data preparation stage. Data preparation, often referred to as “pre-processing” is the stage at which raw data is cleaned up and organized for the following stage of data processing. During preparation, raw data is diligently checked for any errors. The purpose of this step is to eliminate bad data (redundant, incomplete, or incorrect data) and begin to create high-quality data for the best business intelligence.

#### ➤ Data input

The clean data is then entered into its destination (perhaps a CRM like Salesforce or a data warehouse like Redshift), and translated into a language that it can understand. Data input is the first stage in which raw data begins to take the form of usable information.

#### ➤ Processing

During this stage, the data inputted to the computer in the previous stage is actually processed for interpretation. Processing is done using machine learning algorithms, though the process itself may vary slightly depending on the source of data being processed (data lakes, social networks,

connected devices etc.) and its intended use (examining advertising patterns, medical diagnosis from connected devices, determining customer needs, etc.).

➤ Data output/interpretation

The output/interpretation stage is the stage at which data is finally usable to non-data scientists. It is translated, readable, and often in the form of graphs, videos, images, plain text, etc.). Members of the company or institution can now begin to self-serve the data for their own data analytics projects.

➤ Data storage

The final stage of data processing is storage. After all of the data is processed, it is then stored for future use. While some information may be put to use immediately, much of it will serve a purpose later on. Plus, properly stored data is a necessity for compliance with data protection legislation like GDPR. When data is properly stored, it can be quickly and easily accessed by members of the organization when needed.

### **The future of data processing**

The future of data processing lies in the cloud. Cloud technology builds on the convenience of current electronic data processing methods and accelerates its speed and effectiveness. Faster, higher-quality data means more data for each organization to utilize and more valuable insights to extract.

### **2.4. Data Visualization**

Data visualization is the presentation of data in a pictorial or graphical format. It enables decision makers to see analytics presented visually, so they can grasp difficult concepts or identify new patterns. With interactive visualization, you can take the concept a step further by using technology to drill down into charts and graphs for more detail, interactively changing what data you see and how it's processed.

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.

In the world of Big Data, data visualization tools and technologies are essential to analyze massive amounts of information and make data-driven decisions.

## **The different types of visualizations**

When you think of data visualization, your first thought probably immediately goes to simple bar graphs or pie charts. While these may be an integral part of visualizing data and a common baseline for many data graphics, the right visualization must be paired with the right set of information. Simple graphs are only the tip of the iceberg. There's a whole selection of visualization methods to present data in effective and interesting ways.

### **Common general types of data visualization:**

- Charts
- Tables
- Graphs
- Maps
- Infographics
- Dashboards

### **More specific examples of methods to visualize data:**

- Area Chart
- Bar Chart
- Box-and-whisker Plots
- Bubble Cloud
- Bullet Graph
- Cartogram
- Circle View
- Dot Distribution Map
- Gantt Chart
- Heat Map
- Highlight Table
- Histogram
- Matrix
- Network
- Polar Area
- Radial Tree
- Scatter Plot (2D or 3D)
- Streamgraph

- Text Tables
- Timeline
- Treemap
- Wedge Stack Graph
- Word Cloud
- And any mix-and-match combination in a dashboard!

## 2.5 Regression, Prediction & Classification

### Regression Analysis

Regression analysis is a statistical method to model the relationship between a dependent (target) and independent (predictor) variables with one or more independent variables. More specifically, Regression analysis helps us to understand how the value of the dependent variable is changing corresponding to an independent variable when other independent variables are held fixed. It predicts continuous/real values such as temperature, age, salary, price, etc.

We can understand the concept of regression analysis using the below example:

Example: Suppose there is a marketing company A, who does various advertisement every year and get sales on that. The below list shows the advertisement made by the company in the last 5 years and the corresponding sales:

Advertisement	Sales
\$90	\$1000
\$120	\$1300
\$150	\$1800
\$100	\$1200
\$130	\$1380
\$200	??

Table 2.1: Sales data

The company wants to do the advertisement of \$200 in the year 2019 and wants to know the prediction about the sales for this year. So to solve such type of prediction problems in machine learning, we need regression analysis.

Regression is a supervised learning technique which helps in finding the correlation between variables and enables us to predict the continuous output variable based on the one or more predictor variables. It is mainly used for prediction, forecasting, time series modeling, and determining the causal-effect relationship between variables.

In Regression, we plot a graph between the variables which best fits the given datapoints, using this plot, the machine learning model can make predictions about the data. In simple words, **"Regression shows a line or curve that passes through all the datapoints on target-predictor graph in such a way that the vertical distance between the datapoints and the regression line is minimum."** The distance between datapoints and line tells whether a model has captured a strong relationship or not.

Some examples of regression can be as:

- Prediction of rain using temperature and other factors
- Determining Market trends
- Prediction of road accidents due to rash driving

### **Linear Regression:**

- Linear regression is a statistical regression method which is used for predictive analysis.
- It is one of the very simple and easy algorithms which works on regression and shows the relationship between the continuous variables.
- It is used for solving the regression problem in machine learning.
- Linear regression shows the linear relationship between the independent variable (X-axis) and the dependent variable (Y-axis), hence called linear regression.
- If there is only one input variable (x), then such linear regression is called **simple linear regression**. And if there is more than one input variable, then such linear regression is called **multiple linear regression**.
- The relationship between variables in the linear regression model can be explained using the below image. Here we are predicting the salary of an employee on the basis of **the year of experience**.

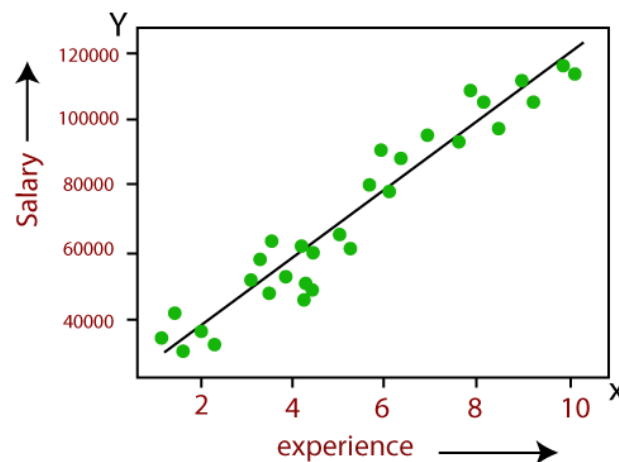


Figure 2.1: Linear regression

- Below is the mathematical equation for Linear regression:

$$Y = aX + b$$

Here, Y = dependent variables (target variables),

X = Independent variables (predictor variables),

a and b are the linear coefficients

Some popular applications of linear regression are:

- Analyzing trends and sales estimates
- Salary forecasting
- Real estate prediction
- Arriving at ETAs in traffic.

### Classification

A classification problem is when the output variable is a category, such as “red” or “blue” or “disease” and “no disease”. A classification model attempts to draw some conclusion from observed values. Given one or more inputs a classification model will try to predict the value of one or more outcomes.

For example, when filtering emails “spam” or “not spam”, when looking at transaction data, “fraudulent”, or “authorized”. In short Classification either predicts categorical class labels or classifies data (construct a model) based on the training set and the values (class labels) in



classifying attributes and uses it in classifying new data. There are a number of classification models. Classification models include logistic regression, decision tree, random forest, gradient-boosted tree, multilayer perceptron, one-vs-rest, and Naive Bayes.

### Types of Classification Algorithms

- Linear Models
  - Logistic Regression
  - Support Vector Machines
- Nonlinear models
  - K-nearest Neighbors (KNN)
  - Kernel Support Vector Machines (SVM)
  - Naïve Bayes
  - Decision Tree Classification
  - Random Forest Classification

## 2.6. Clustering & Recommender Systems

### Introduction to Clustering

It is basically a type of unsupervised learning method . An unsupervised learning method is a method in which we draw references from datasets consisting of input data without labelled responses. Generally, it is used as a process to find meaningful structure, explanatory underlying processes, generative features, and groupings inherent in a set of examples.

**Clustering** is the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar to other data points in the same group and dissimilar to the data points in other groups. It is basically a collection of objects on the basis of similarity and dissimilarity between them.

**For ex–** The data points in the graph below clustered together can be classified into one single group. We can distinguish the clusters, and we can identify that there are 3 clusters in the below picture.

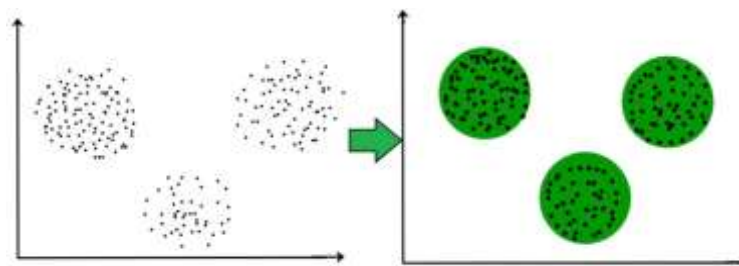


Figure 2.2: Clustering

**Clustering Methods :**

- **Density-Based Methods :** These methods consider the clusters as the dense region having some similarity and different from the lower dense region of the space. These methods have good accuracy and ability to merge two clusters. Example DBSCAN (Density-Based Spatial Clustering of Applications with Noise), OPTICS (Ordering Points to Identify Clustering Structure) etc.
- **Hierarchical Based Methods :** The clusters formed in this method forms a tree-type structure based on the hierarchy. New clusters are formed using the previously formed one. It is divided into two category
  - **Agglomerative** (bottom up approach)
  - **Divisive** (top down approach)examples CURE (Clustering Using Representatives), BIRCH (Balanced Iterative Reducing Clustering and using Hierarchies) etc.
- **Partitioning Methods :** These methods partition the objects into k clusters and each partition forms one cluster. This method is used to optimize an objective criterion similarity function such as when the distance is a major parameter example K-means, CLARANS (Clustering Large Applications based upon Randomized Search) etc.
- **Grid-based Methods :** In this method the data space is formulated into a finite number of cells that form a grid-like structure. All the clustering operation done on these grids are fast and independent of the number of data objects example STING (Statistical Information Grid), wave cluster, CLIQUE (CLustering In Quest) etc.

**Recommender systems**

Recommender systems are the systems that are designed to recommend things to the user based on many different factors. These systems predict the most likely product that the users are most likely to purchase and are of interest to. Companies like Netflix, Amazon, etc. use recommender systems to help their users to identify the correct product or movies for them.

The recommender system deals with a large volume of information present by filtering the most important information based on the data provided by a user and other factors that take care of the user's preference and interest. It finds out the match between user and item and imputes the similarities between users and items for recommendation.

Both the users and the services provided have benefited from these kinds of systems. The quality and decision-making process has also improved through these kinds of systems.

Why the Recommendation system?

- Benefits users in finding items of their interest.
- Help item providers in delivering their items to the right user.
- Identity products that are most relevant to users.
- Personalized content.
- Help websites to improve user engagement.

What can be Recommended?

There are many different things that can be recommended by the system like movies, books, news, articles, jobs, advertisements, etc. Netflix uses a recommender system to recommend movies & web-series to its users. Similarly, YouTube recommends different videos. There are many examples of recommender systems that are widely used today.

### **Types of Recommendation System**

#### ➤ Popularity-Based Recommendation System

It is a type of recommendation system which works on the principle of popularity and or anything which is in trend. These systems check about the product or movie which are in trend or are most popular among the users and directly recommend those.

For example, if a product is often purchased by most people then the system will get to know that that product is most popular so for every new user who just signed it, the system will recommend that product to that user also and chances becomes high that the new user will also purchase that.

#### ➤ Merits of popularity based recommendation system

- It does not suffer from cold start problems which means on day 1 of the business also it can recommend products on various different filters.
- There is no need for the user's historical data.
  - Demerits of popularity based recommendation system
- Not personalized
- The system would recommend the same sort of products/movies which are solely based upon popularity to every other user.

**Example**

- Google News: News filtered by trending and most popular news.
- YouTube: Trending videos.

**Questions**

1. What is Data and Big Data?
2. What is algorithm and its properties?
3. Explain data and its acquisition.
4. What are the stages involved in data processing?
5. Define data visualization.
6. How many types of data visualization.
7. What is data classification and Regression?
8. What is data clustering? Explain any one method in details.
9. What are recommender systems? How is working in OTT.
10. How many types of data acquisition systems.

### Unit-III

#### Natural Language Processing

Natural Language Processing (NLP) refers to AI method of communicating with an intelligent systems using a natural language such as English. Processing of Natural Language is required when you want an intelligent system like robot to perform as per your instructions, when you want to hear decision from a dialogue based clinical expert system, etc.

The field of NLP involves making computers to perform useful tasks with the natural languages humans use. The input and output of an NLP system can be –

- Speech
- Written Text

#### Components of NLP

There are two components of NLP as given

##### ➤ Natural Language Understanding (NLU)

Understanding involves the following tasks –

- Mapping the given input in natural language into useful representations.
- Analyzing different aspects of the language.

##### ➤ Natural Language Generation (NLG)

It is the process of producing meaningful phrases and sentences in the form of natural language from some internal representation.

It involves –

- **Text planning** – It includes retrieving the relevant content from knowledge base.
- **Sentence planning** – It includes choosing required words, forming meaningful phrases, setting tone of the sentence.
- **Text Realization** – It is mapping sentence plan into sentence structure.

The NLU is harder than NLG.

#### Difficulties in NLU

NL has an extremely rich form and structure.

It is very ambiguous. There can be different levels of ambiguity –

- **Lexical ambiguity** – It is at very primitive level such as word-level.
- For example, treating the word “board” as noun or verb?
- **Syntax Level ambiguity** – A sentence can be parsed in different ways.
- For example, “He lifted the beetle with red cap.” – Did he use cap to lift the beetle or he lifted a beetle that had red cap?
- **Referential ambiguity** – Referring to something using pronouns. For example, Rima went to Gauri. She said, “I am tired.” – Exactly who is tired?
- One input can mean different meanings.
- Many inputs can mean the same thing.

### NLP Terminology

- **Phonology** – It is study of organizing sound systematically.
- **Morphology** – It is a study of construction of words from primitive meaningful units.
- **Morpheme** – It is primitive unit of meaning in a language.
- **Syntax** – It refers to arranging words to make a sentence. It also involves determining the structural role of words in the sentence and in phrases.
- **Semantics** – It is concerned with the meaning of words and how to combine words into meaningful phrases and sentences.
- **Pragmatics** – It deals with using and understanding sentences in different situations and how the interpretation of the sentence is affected.
- **Discourse** – It deals with how the immediately preceding sentence can affect the interpretation of the next sentence.
- **World Knowledge** – It includes the general knowledge about the world.

### Steps in NLP

There are general five steps –

- **Lexical Analysis** – It involves identifying and analyzing the structure of words. Lexicon of a language means the collection of words and phrases in a language. Lexical analysis is dividing the whole chunk of txt into paragraphs, sentences, and words.

- **Syntactic Analysis (Parsing)** – It involves analysis of words in the sentence for grammar and arranging words in a manner that shows the relationship among the words. The sentence such as “The school goes to boy” is rejected by English syntactic analyzer.

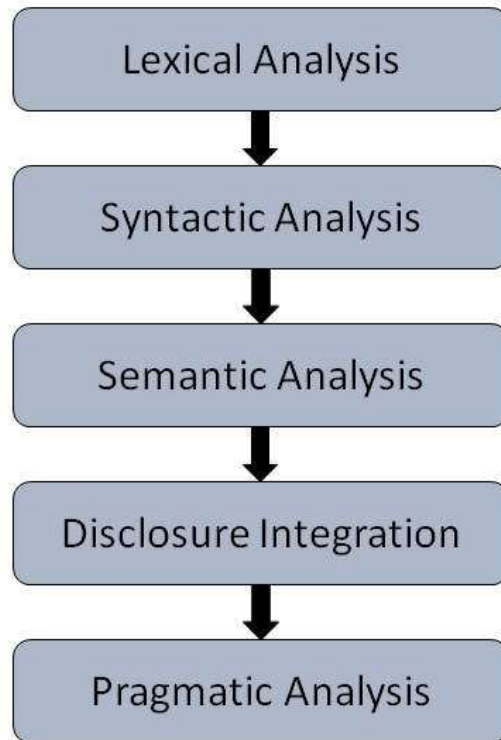


Figure 3.1: NLP Steps

- **Semantic Analysis** – It draws the exact meaning or the dictionary meaning from the text. The text is checked for meaningfulness. It is done by mapping syntactic structures and objects in the task domain. The semantic analyzer disregards sentence such as “hot ice-cream”.
- **Discourse Integration** – The meaning of any sentence depends upon the meaning of the sentence just before it. In addition, it also brings about the meaning of immediately succeeding sentence.
- **Pragmatic Analysis** – During this, what was said is re-interpreted on what it actually meant. It involves deriving those aspects of language which require real world knowledge.

### 3.1 Speech Recognition

Speech recognition, also known as automatic speech recognition (ASR), computer speech recognition, or speech-to-text, is a capability which enables a program to process human speech into a written format. While it's commonly confused with voice recognition, speech recognition

focuses on the translation of speech from a verbal format to a text one whereas voice recognition just seeks to identify an individual user's voice.

Many speech recognition applications and devices are available, but the more advanced solutions use AI and machine learning. They integrate grammar, syntax, structure, and composition of audio and voice signals to understand and process human speech. Ideally, they learn as they go — evolving responses with each interaction.

The best kind of systems also allow organizations to customize and adapt the technology to their specific requirements — everything from language and nuances of speech to brand recognition. For example:

- **Language weighting:** Improve precision by weighting specific words that are spoken frequently (such as product names or industry jargon), beyond terms already in the base vocabulary.
- **Speaker labeling:** Output a transcription that cites or tags each speaker's contributions to a multi-participant conversation.
- **Acoustics training:** Attend to the acoustical side of the business. Train the system to adapt to an acoustic environment (like the ambient noise in a call center) and speaker styles (like voice pitch, volume and pace).
- **Profanity filtering:** Use filters to identify certain words or phrases and sanitize speech output.

Meanwhile, speech recognition continues to advance. Companies, like IBM, are making inroads in several areas, the better to improve human and machine interaction.

### Speech recognition algorithms

The vagaries of human speech have made development challenging. It's considered to be one of the most complex areas of computer science – involving linguistics, mathematics and statistics. Speech recognizers are made up of a few components, such as the speech input, feature extraction, feature vectors, a decoder, and a word output. The decoder leverages acoustic models, a pronunciation dictionary, and language models to determine the appropriate output.



Speech recognition technology is evaluated on its accuracy rate, i.e. word error rate (WER), and speed. A number of factors can impact word error rate, such as pronunciation, accent, pitch, volume, and background noise. Reaching human parity – meaning an error rate on par with that of two humans speaking – has long been the goal of speech recognition systems. Research from Lippmann (link resides outside IBM) estimates the word error rate to be around 4 percent, but it's been difficult to replicate the results from this paper.

Read more on how IBM has made strides in this respect, achieving industry records in the field of speech recognition.

Various algorithms and computation techniques are used to recognize speech into text and improve the accuracy of transcription. Below are brief explanations of some of the most commonly used methods:

- **Natural language processing (NLP):** While NLP isn't necessarily a specific algorithm used in speech recognition, it is the area of artificial intelligence which focuses on the interaction between humans and machines through language through speech and text. Many mobile devices incorporate speech recognition into their systems to conduct voice search—e.g. Siri—or provide more accessibility around texting.
- **Hidden markov models (HMM):** Hidden Markov Models build on the Markov chain model, which stipulates that the probability of a given state hinges on the current state, not its prior states. While a Markov chain model is useful for observable events, such as text inputs, hidden markov models allow us to incorporate hidden events, such as part-of-speech tags, into a probabilistic model. They are utilized as sequence models within speech recognition, assigning labels to each unit—i.e. words, syllables, sentences, etc.—in the sequence. These labels create a mapping with the provided input, allowing it to determine the most appropriate label sequence.
- **N-grams:** This is the simplest type of language model (LM), which assigns probabilities to sentences or phrases. An N-gram is sequence of N-words. For example, “order the pizza” is a trigram or 3-gram and “please order the pizza” is a 4-gram. Grammar and the probability of certain word sequences are used to improve recognition and accuracy.
- **Neural networks:** Primarily leveraged for deep learning algorithms, neural networks process training data by mimicking the interconnectivity of the human brain through layers of nodes. Each node is made up of inputs, weights, a bias (or threshold) and an output. If

that output value exceeds a given threshold, it “fires” or activates the node, passing data to the next layer in the network. Neural networks learn this mapping function through supervised learning, adjusting based on the loss function through the process of gradient descent. While neural networks tend to be more accurate and can accept more data, this comes at a performance efficiency cost as they tend to be slower to train compared to traditional language models.

- **Speaker Diarization (SD):** Speaker diarization algorithms identify and segment speech by speaker identity. This helps programs better distinguish individuals in a conversation and is frequently applied at call centers distinguishing customers and sales agents.

### 3.2. Natural language understanding

Natural language understanding (NLU) is a sub-topic of natural language processing, which involves breaking down the human language into a machine-readable format. Interesting applications include text categorization, machine translation, and question answering.

NLU uses grammatical rules and common syntax to understand the overall context and meaning of “natural language,” beyond literal definitions. Its goal is to understand written or spoken language the way a human would.

NLU is used in natural language processing (NLP) tasks like topic classification, language detection, and sentiment analysis:

- Sentiment analysis automatically interprets emotions within a text and categorizes them as positive, negative, or neutral. By quickly understanding, processing, and analyzing thousands of online conversations, sentiment analysis tools can deliver valuable insights about how customers view your brand and products.

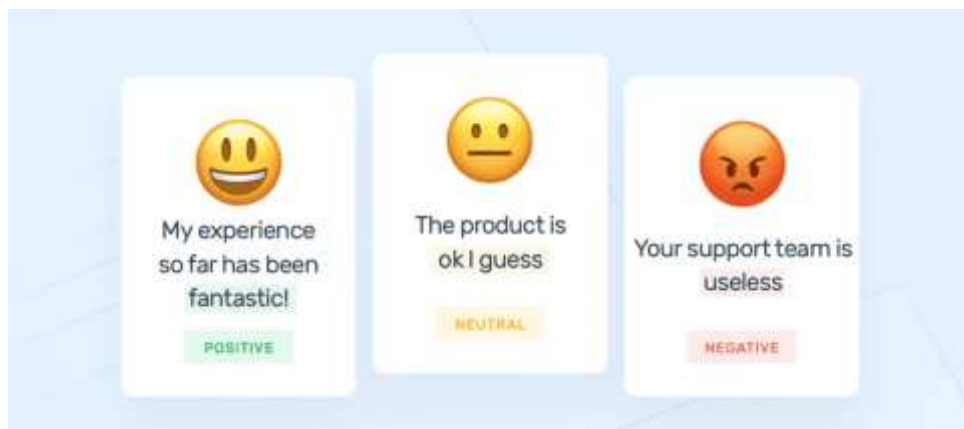


Figure 3.2: Natural language understanding

- Language detection automatically understands the language of written text. An essential tool to help businesses route tickets to the correct local teams, avoid wasting time passing tickets from one customer agent to the next, and respond to customer issues faster.
- Topic classification is able to understand natural language to automatically sort texts into predefined groups or topics. Software company Atlassian, for example, uses the tags Reliability, Usability, and Functionality to sort incoming customer support tickets, enabling them to deal with customer issues efficiently.

### NLP Vs NLU: What's The Difference?

Natural language understanding is a subfield of natural language processing.

While both NLP and NLU aim to make sense of unstructured data, but they are not the same thing. NLP is concerned with how computers are programmed to process language and facilitate “natural” back-and-forth communication between computers and humans. Natural language understanding, on the other hand, focuses on a machine’s ability to understand the human language. NLU refers to how unstructured data is rearranged so that machines may “understand” and analyze it.

Look at it this way. Before a computer can **process** unstructured text into a machine-readable format, first machines need to **understand** the peculiarities of the human language.

### Natural Language Understanding Examples

#### ➤ Machine Translation (MT)

Accurately translating text or speech from one language to another is one of the toughest challenges of natural language processing and natural language understanding.

Using complex algorithms that rely on linguistic rules and AI machine training, Google Translate, Microsoft Translator, and Facebook Translation have become leaders in the field of “generic” language translation.

You can type text or upload whole documents and receive translations in dozens of languages using machine translation tools. Google Translate even includes optical character recognition (OCR) software, which allows machines to extract text from images, read and translate it.

#### ➤ Automated Reasoning

Automated reasoning is a subfield of cognitive science that is used to automatically prove mathematical theorems or make logical inferences about a medical diagnosis. It gives machines a form of reasoning or logic, and allows them to infer new facts by deduction.

Simply put, using previously gathered and analyzed information, computer programs are able to generate conclusions. For example, in medicine, machines can infer a diagnosis based on previous diagnoses using IF-THEN deduction rules.

#### ➤ Automatic Routing of Tickets

A useful business example of NLU is customer service automation. With text analysis solutions like MonkeyLearn, machines can understand the content of customer support tickets and route them to the correct departments without employees having to open every single ticket. Not only does this save customer support teams hundreds of hours, but it also helps them prioritize urgent tickets.

According to Zendesk, tech companies receive more than 2,600 customer support inquiries per month. Using NLU technology, you can sort unstructured data (email, social media, live chat, etc.) by topic, sentiment, and urgency (among others). These tickets can then be routed directly to the relevant agent and prioritized.

#### ➤ Question Answering

Question answering is a subfield of NLP and speech recognition that uses NLU to help computers automatically understand natural language questions. For example, here's a common question you might ask Google Assistant: "What's the weather like tomorrow?" NLP tools can split this question into topic (weather) and date (tomorrow), understand it and gather the most appropriate answer from unstructured collections of "natural language documents": online news reports, collected web pages, reference texts, etc

By default, virtual assistants tell you the weather for your current location, unless you specify a particular city. The goal of question answering is to give the user response in their natural language, rather than a list of text answers.

### 3.3. Natural language generation

Natural-language generation (NLG) is a software process that transforms structured data into natural language. It can be used to produce long form content for organizations to automate custom reports, as well as produce custom content for a web or mobile application. It can also be used to generate short blurbs of text in interactive conversations (a chatbot) which might even be read out by a text-to-speech system.

Automated NLG can be compared to the process humans use when they turn ideas into writing or speech. Psycholinguists prefer the term language production for this process, which can also be described in mathematical terms, or modeled in a computer for psychological research. NLG systems can also be compared to translators of artificial computer languages, such as decompilers or transpilers, which also produce human-readable code generated from an intermediate representation. Human languages tend to be considerably more complex and allow for much more ambiguity and variety of expression than programming languages, which makes NLG more challenging.

NLG may be viewed as the opposite of natural-language understanding (NLU): whereas in natural-language understanding, the system needs to disambiguate the input sentence to produce the machine representation language, in NLG the system needs to make decisions about how to put a concept into words. The practical considerations in building NLU vs. NLG systems are not symmetrical. NLU needs to deal with ambiguous or erroneous user input, whereas the ideas the system wants to express through NLG are generally known precisely. NLG needs to choose a specific, self-consistent textual representation from many potential representations, whereas NLU generally tries to produce a single, normalized representation of the idea expressed.

### 3.4. Chatbots

A chatbot is an artificial intelligence (AI) software that can simulate a conversation (or a chat) with a user in natural language through messaging applications, websites, mobile apps or through the telephone.

Why are chatbots important? A chatbot is often described as one of the most advanced and promising expressions of interaction between humans and machines. However, from a technological point of view, a chatbot only represents the natural evolution of a Question

Answering system leveraging Natural Language Processing (NLP). Formulating responses to questions in natural language is one of the most typical Examples of Natural Language Processing applied in various enterprises' end-use applications.

Behind the scenes: How a chatbot works

There are two different tasks at the core of a chatbot:

- 1) User request analysis
- 2) Returning the response



1) User request analysis: this is the first task that a chatbot performs. It analyzes the user's request to identify the user intent and to extract relevant entities.

The ability to identify the user's intent and extract data and relevant entities contained in the user's request is the first condition and the most relevant step at the core of a chatbot: If you are not able to correctly understand the user's request, you won't be able to provide the correct answer.



2) Returning the response: once the user's intent has been identified, the chatbot must provide the most appropriate response for the user's request. The answer may be:

- a generic and predefined text
- a text retrieved from a knowledge base that contains different answers
- a contextualized piece of information based on data the user has provided
- data stored in enterprise systems

- the result of an action that the chatbot performed by interacting with one or more backend application
- a disambiguating question that helps the chatbot to correctly understand the user's request

### **Why chatbots are important**

Chatbot applications streamline interactions between people and services, enhancing customer experience. At the same time, they offer companies new opportunities to improve the customers engagement process and operational efficiency by reducing the typical cost of customer service.

To be successful, a chatbot solution should be able to effectively perform both of these tasks. Human support plays a key role here: Regardless of the kind of approach and the platform, human intervention is crucial in configuring, training and optimizing the chatbot system.

### **3.5. Machine Translation**

Machine translation (MT) refers to fully automated software that can translate source content into target languages. Humans may use MT to help them render text and speech into another language, or the MT software may operate without human intervention.

MT tools are often used to translate vast amounts of information involving millions of words that could not possibly be translated the traditional way. The quality of MT output can vary considerably; MT systems require “training” in the desired domain and language pair to increase quality.

Translation companies use MT to augment productivity of their translators, cut costs, and provide post-editing services to clients. MT use by language service providers is growing quickly. In 2016, SDL—one of the largest translation companies in the world—announced it translates 20 times more content with MT than with human teams.

Customizable MT refers to MT software that has a basic component and can be trained to improve terminology accuracy in a chosen domain (medical, legal, IP, or a company's own preferred terminology). For example, WIPO's specialist MT engine translates patents more accurately than generalist MT engines, and eBay's solution can understand and render into other languages hundreds of abbreviations used in electronic commerce.

Adaptive MT offers suggestions to translators as they type in their CAT-tool, and learns from their input continuously in real time. Introduced by Lilt in 2016 and by SDL in 2017, adaptive MT is

believed to improve translator productivity significantly and can challenge translation memory technology in the future.

There are over 100 providers of MT technologies. Some of them are strictly MT developers, others are translation firms and IT giants.

### Examples of MT Providers

Google Translate	Microsoft Translator / Bing	SDL BeGlobal
Yandex Translate	Amazon Web Services translator	Naver
IBM - Watson Language Translator	Automatic Trans	BABYLON
CCID TransTech Co.	CSLi	East Linden
Eleka Ingeniaritza Linguistikoa	GrammarSoft ApS	Iconic Translation Machines
K2E-PAT	KantanMT	Kodensha
Language Engineering Company	Lighthouse IP Group	Lingenio
Lingosail Technology Co.	LionBridge	Lucy Software / ULG
MorphoLogic / Globalese	Multilizer	NICT
Omniscien	Pangeanic	Precision Translation Tools (Slate)
Prompsit Language Engineering	PROMT	Raytheon
Reverso Softissimo	SkyCode	Smart Communications
Sovee	SyNTHEMA	SYSTRAN
tauyou	Tilde	Trident Software

Table 3.1: MT providers



## MT Approaches

There are three main approaches to machine translation:

- First-generation rule-based (RbMT) systems rely on countless algorithms based on the grammar, syntax, and phraseology of a language.
- Statistical systems (SMT) arrived with search and big data. With lots of parallel texts becoming available, SMT developers learned to pattern-match reference texts to find translations that are statistically most likely to be suitable. These systems train faster than RbMT, provided there is enough existing language material to reference.
- Neural MT (NMT) uses machine learning technology to teach software how to produce the best result. This process consumes large amounts of processing power, and that is why it's often run on graphics units of CPUs. NMT started gaining visibility in 2016. Many MT providers are now switching to this technology.

A combination of two different MT methods is called Hybrid MT.

## Questions

1. What is Natural Language Processing? Discuss with some applications.
2. List any two real-life applications of Natural Language Processing.
3. What is Speech recognition
4. Explain the Natural language understanding and Natural language generation
5. Show the working of chatbots.
6. Analyse how statistical methods can be used in machine translation
7. Describe the different components of a typical conversational agent

## Unit-4

### Artificial Neural Networks

Artificial intelligence (AI), also known as machine intelligence, is a branch of computer science that aims to imbue software with the ability to analyze its environment using either predetermined rules and search algorithms, or pattern recognizing machine learning models, and then make decisions based on those analyses.

#### Basic Structure of ANNs

The idea of ANNs is based on the belief that working of human brain by making the right connections, can be imitated using silicon and wires as living neurons and dendrites.

The human brain is composed of 86 billion nerve cells called neurons. They are connected to other thousand cells by Axons. Stimuli from external environment or inputs from sensory organs are accepted by dendrites. These inputs create electric impulses, which quickly travel through the neural network. A neuron can then send the message to other neuron to handle the issue or does not send it forward.

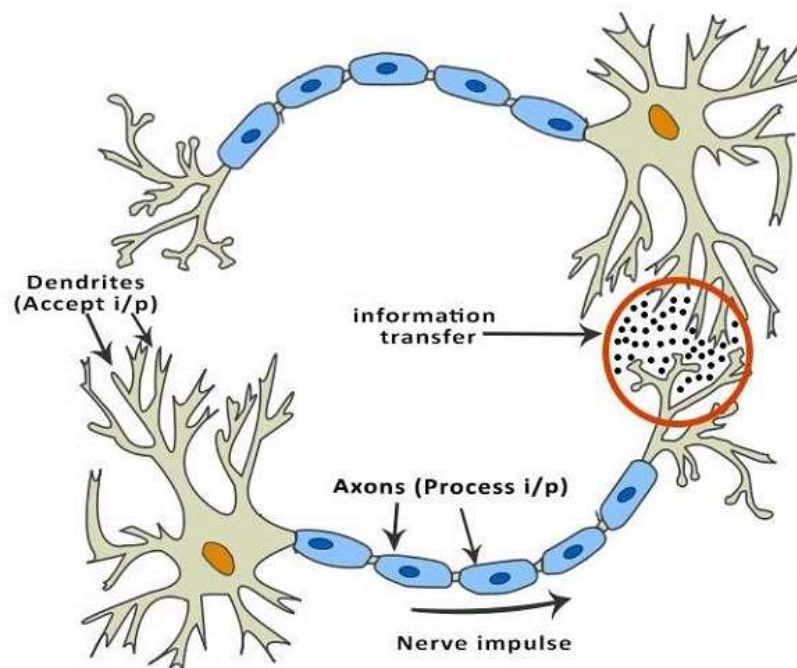


Figure 4.1: Neuron

ANNs are composed of multiple **nodes**, which imitate biological **neurons** of human brain. The neurons are connected by links and they interact with each other. The nodes can take input data and perform simple operations on the data. The result of these operations is passed to other neurons. The output at each node is called its **activation** or **node value**.

Each link is associated with **weight**. ANNs are capable of learning, which takes place by altering weight values. The following illustration shows a simple ANN –

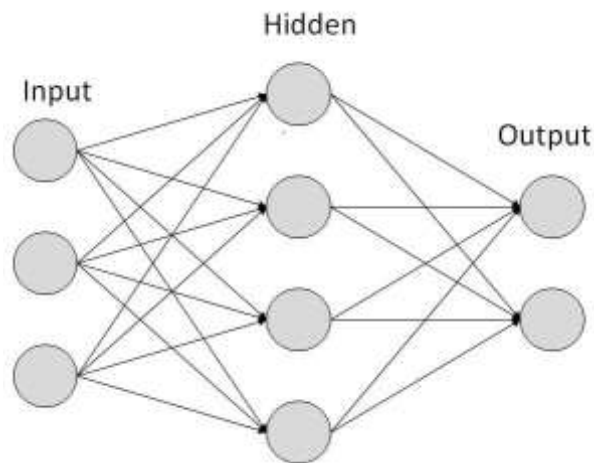


Figure 4.2: ANN

### Types of Artificial Neural Networks

There are two Artificial Neural Network topologies – **FeedForward** and **Feedback**.

#### ➤ FeedForward ANN

In this ANN, the information flow is unidirectional. A unit sends information to other unit from which it does not receive any information. There are no feedback loops. They are used in pattern generation/recognition/classification. They have fixed inputs and outputs.

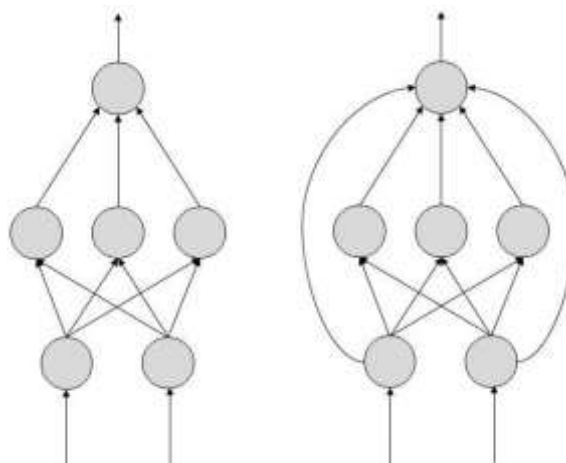


Figure 4.2: FeedForward ANN

#### ➤ FeedBack ANN

Here, feedback loops are allowed. They are used in content addressable memories.

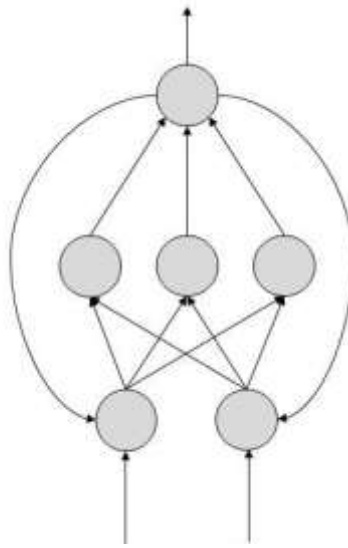


Figure 4.3: FeedBack ANN

#### 4.1. Deep Learning

Deep learning is a branch of machine learning which is completely based on artificial neural networks, as neural network is going to mimic the human brain so deep learning is also a kind of mimic of human brain. In deep learning, we don't need to explicitly program everything. The concept of deep learning is not new. It has been around for a couple of years now. It's on hype nowadays because earlier we did not have that much processing power and a lot of data

##### Architectures :

1. **Deep Neural Network** – It is a neural network with a certain level of complexity (having multiple hidden layers in between input and output layers). They are capable of modeling and processing non-linear relationships.
2. **Deep Belief Network(DBN)** – It is a class of Deep Neural Network. It is multi-layer belief networks.

##### Steps for performing DBN :

- a. Learn a layer of features from visible units using Contrastive Divergence algorithm.
  - b. Treat activations of previously trained features as visible units and then learn features of features.
  - c. Finally, the whole DBN is trained when the learning for the final hidden layer is achieved.
3. **Recurrent** (perform same task for every element of a sequence) **Neural Network** – Allows for parallel and sequential computation. Similar to the human brain (large feedback network of connected neurons). They are able to remember important things about the input they received and hence enables them to be more precise.

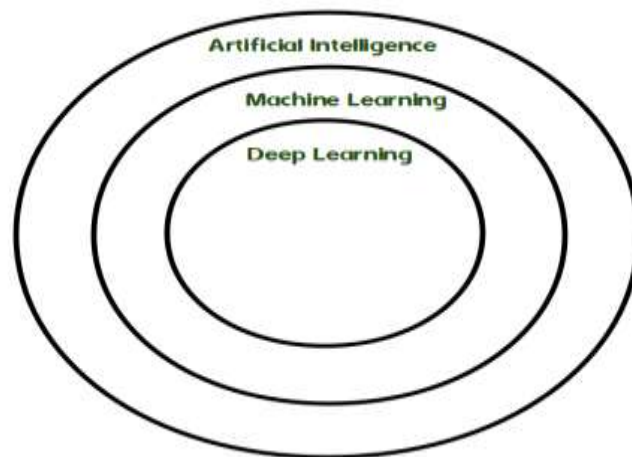


Figure 4.4: Deep Learning

**Working:**

First, we need to identify the actual problem in order to get the right solution and it should be understood, the feasibility of the Deep Learning should also be checked (whether it should fit Deep Learning or not). Second, we need to identify the relevant data which should correspond to the actual problem and should be prepared accordingly. Third, Choose the Deep Learning Algorithm appropriately. Fourth, Algorithm should be used while training the dataset. Fifth, Final testing should be done on the dataset.

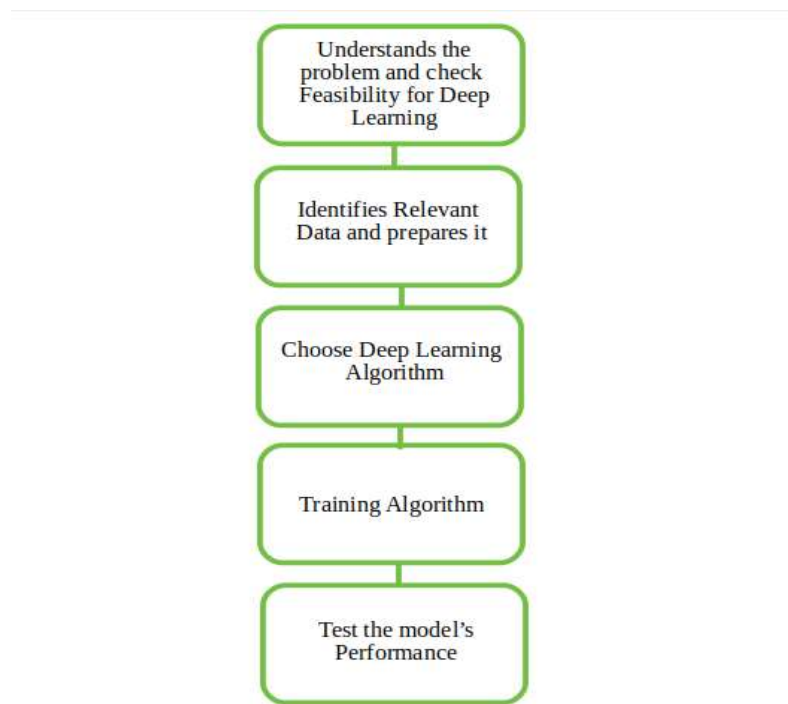


Figure 4.5: Deep Learning Working

## 4.2. Recurrent Neural Networks

**Recurrent Neural Network(RNN)** are a type of Neural Network where the **output from previous step are fed as input to the current step**. In traditional neural networks, all the inputs and outputs are independent of each other, but in cases like when it is required to predict the next word of a sentence, the previous words are required and hence there is a need to remember the previous words. Thus RNN came into existence, which solved this issue with the help of a Hidden Layer. The main and most important feature of RNN is **Hidden state**, which remembers some information about a sequence.

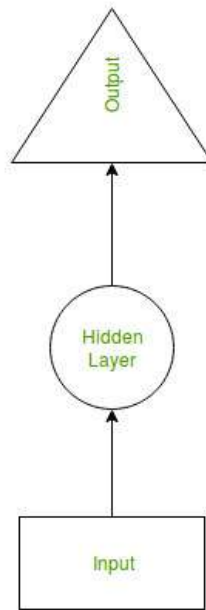


Figure 4.6: Recurrent Neural Networks

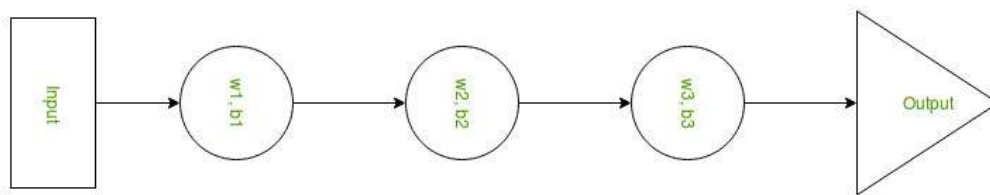
RNN have a “**memory**” which remembers all information about what has been calculated. It uses the same parameters for each input as it performs the same task on all the inputs or hidden layers to produce the output. This reduces the complexity of parameters, unlike other neural networks.

### How RNN works

The working of a RNN can be understood with the help of below example:

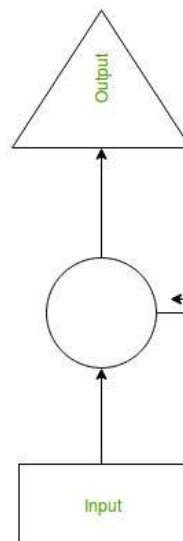
#### Example:

Suppose there is a deeper network with one input layer, three hidden layers and one output layer. Then like other neural networks, each hidden layer will have its own set of weights and biases, let's say, for hidden layer 1 the weights and biases are  $(w_1, b_1)$ ,  $(w_2, b_2)$  for second hidden layer and  $(w_3, b_3)$  for third hidden layer. This means that each of these layers are independent of each other, i.e. they do not memorize the previous outputs.



Now the RNN will do the following:

- RNN converts the independent activations into dependent activations by providing the same weights and biases to all the layers, thus reducing the complexity of increasing parameters and memorizing each previous outputs by giving each output as input to the next hidden layer.
- Hence these three layers can be joined together such that the weights and bias of all the hidden layers is the same, into a single recurrent layer.



- **Formula for calculating current state:**

$$h_t = f(h_{t-1}, x_t)$$

where:

$h_t$  -> current state

$h_{t-1}$  -> previous state

$x_t$  -> input state

- **Formula for applying Activation function(tanh):**

$$h_t = \tanh (W_{hh}h_{t-1} + W_{xh}x_t)$$

where:

$W_{hh}$  -> weight at recurrent neuron

$W_{xh}$  -> weight at input neuron

- **Formula for calculating output:**

$$y_t = W_{hy}h_t$$

$Y_t$  -> output

$W_{hy}$  -> weight at output layer



### 4.3. Convolutional Neural Networks

Convolutional Neural Networks are very similar to ordinary Neural Networks from the previous chapter: they are made up of neurons that have learnable weights and biases. Each neuron receives some inputs, performs a dot product and optionally follows it with a non-linearity. The whole network still expresses a single differentiable score function: from the raw image pixels on one end to class scores at the other. And they still have a loss function (e.g. SVM/Softmax) on the last (fully-connected) layer and all the tips/tricks we developed for learning regular Neural Networks still apply.

#### Architecture Overview

Recall: Regular Neural Nets. As we saw in the previous chapter, Neural Networks receive an input (a single vector), and transform it through a series of hidden layers. Each hidden layer is made up of a set of neurons, where each neuron is fully connected to all neurons in the previous layer, and where neurons in a single layer function completely independently and do not share any connections. The last fully-connected layer is called the “output layer” and in classification settings it represents the class scores.

Regular Neural Nets don’t scale well to full images. In CIFAR-10, images are only of size  $32 \times 32 \times 3$  (32 wide, 32 high, 3 color channels), so a single fully-connected neuron in a first hidden layer of a regular Neural Network would have  $32 \times 32 \times 3 = 3072$  weights. This amount still seems manageable, but clearly this fully-connected structure does not scale to larger images. For example, an image of more respectable size, e.g.  $200 \times 200 \times 3$ , would lead to neurons that have  $200 \times 200 \times 3 = 120,000$  weights. Moreover, we would almost certainly want to have several such neurons, so the parameters would add up quickly! Clearly, this full connectivity is wasteful and the huge number of parameters would quickly lead to overfitting.

#### 3D volumes of neurons.

Convolutional Neural Networks take advantage of the fact that the input consists of images and they constrain the architecture in a more sensible way. In particular, unlike a regular Neural Network, the layers of a ConvNet have neurons arranged in 3 dimensions: width, height, depth. (Note that the word depth here refers to the third dimension of an activation volume, not to the depth of a full Neural Network, which can refer to the total number of layers in a network.) For example, the input images in CIFAR-10 are an input volume of activations, and the volume has dimensions  $32 \times 32 \times 3$  (width, height, depth respectively). As we will soon see, the neurons in a layer will only be connected to a small region of the layer before it, instead of all of the neurons in a fully-

connected manner. Moreover, the final output layer would for CIFAR-10 have dimensions  $1 \times 1 \times 10$ , because by the end of the ConvNet architecture we will reduce the full image into a single vector of class scores, arranged along the depth dimension. Here is a visualization

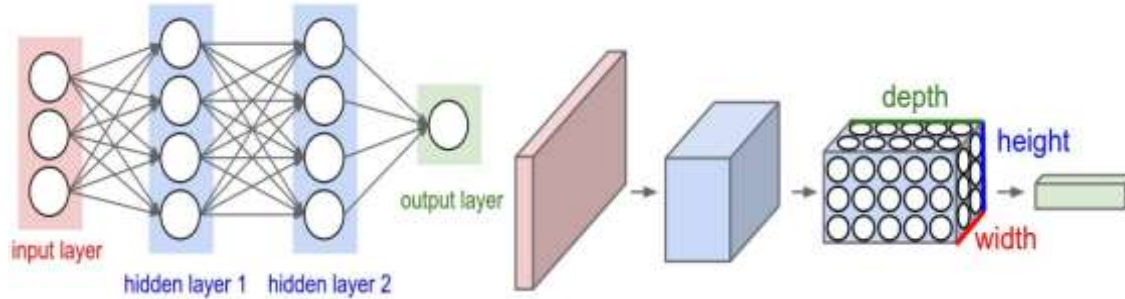


Figure 4.7: Convolutional Neural Networks

#### 4.4. The Universal Approximation Theorem

Mathematically speaking, any neural network architecture aims at finding any mathematical function  $y = f(x)$  that can map attributes( $x$ ) to output( $y$ ). The accuracy of this function i.e. mapping differs depending on the distribution of the dataset and the architecture of the network employed. The function  $f(x)$  can be arbitrarily complex. The Universal Approximation Theorem tells us that Neural Networks has a kind of **universality** i.e. no matter what  $f(x)$  is, there is a network that can approximately approach the result and do the job! This result holds for any number of inputs and outputs.

If we observe the neural network above, considering the input attributes provided as height and width, our job is to predict the gender of the person. If we exclude all the activation layers from the above network, we realize that  $h_1$  is a linear function of both weight and height with parameters  $w_1$ ,  $w_2$ , and the bias term  $b_1$ . Therefore mathematically,

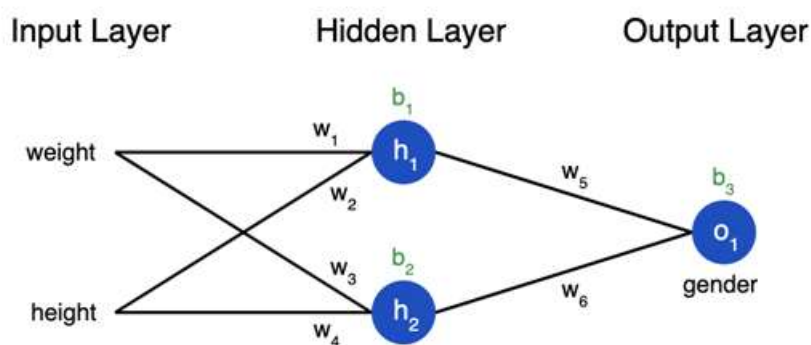


Figure 4.8: Universal Approximation

$$h_1 = w_1 * \text{weight} + w_2 * \text{height} + b_1$$

Similarly,

$$h_2 = w_3 * \text{weight} + w_4 * \text{height} + b_2$$

Going along these lines we realize that  $o_1$  is also a linear function of  $h_1$  and  $h_2$ , and therefore depends linearly on input attributes weight and height as well. This essentially boils down to a linear regression model. Does a linear function suffice at approaching the Universal Approximation Theorem? The answer is NO. This is where activation layers come into play. An activation layer is applied right after a linear layer in the Neural Network to provide non-linearities. Non-linearities help Neural Networks perform more complex tasks. An activation layer operates on activations ( $h_1$ ,  $h_2$  in this case) and modifies them according to the activation function provided for that particular activation layer. Activation functions are generally non-linear except for the identity function. Some commonly used activation functions are ReLu, sigmoid, softmax, etc. With the introduction of non-linearity's along with linear terms, it becomes possible for a neural network to model any given function approximately on having appropriate parameters( $w_1$ ,  $w_2$ ,  $b_1$ , etc in this case). The parameters converge to appropriateness on training suitably. You can get better acquainted mathematically with the Universal Approximation theorem from here.

#### 4.5. Generative Adversarial Networks

Generative Adversarial Networks (GANs) are a powerful class of neural networks that are used for unsupervised learning. It was developed and introduced by Ian J. Goodfellow in 2014. GANs are basically made up of a system of two competing neural network models which compete with each other and are able to analyze, capture and copy the variations within a dataset.

##### Why were GANs developed in the first place?

It has been noticed most of the mainstream neural nets can be easily fooled into misclassifying things by adding only a small amount of noise into the original data. Surprisingly, the model after adding noise has higher confidence in the wrong prediction than when it predicted correctly. The reason for such adversary is that most machine learning models learn from a limited amount of data, which is a huge drawback, as it is prone to overfitting. Also, the mapping between the input and the output is almost linear. Although, it may seem that the boundaries of separation between the various classes are linear, but in reality, they are composed of linearities and even a small change in a point in the feature space might lead to misclassification of data.

## How does GANs work?

Generative Adversarial Networks (GANs) can be broken down into three parts:

- **Generative:** To learn a generative model, which describes how data is generated in terms of a probabilistic model.
- **Adversarial:** The training of a model is done in an adversarial setting.
- **Networks:** Use deep neural networks as the artificial intelligence (AI) algorithms for training purpose.

In GANs, there is a **generator** and a **discriminator**. The Generator generates fake samples of data (be it an image, audio, etc.) and tries to fool the Discriminator. The Discriminator, on the other hand, tries to distinguish between the real and fake samples. The Generator and the Discriminator are both Neural Networks and they both run in competition with each other in the training phase. The steps are repeated several times and in this, the Generator and Discriminator get better and better in their respective jobs after each repetition. The working can be visualized by the diagram given below:

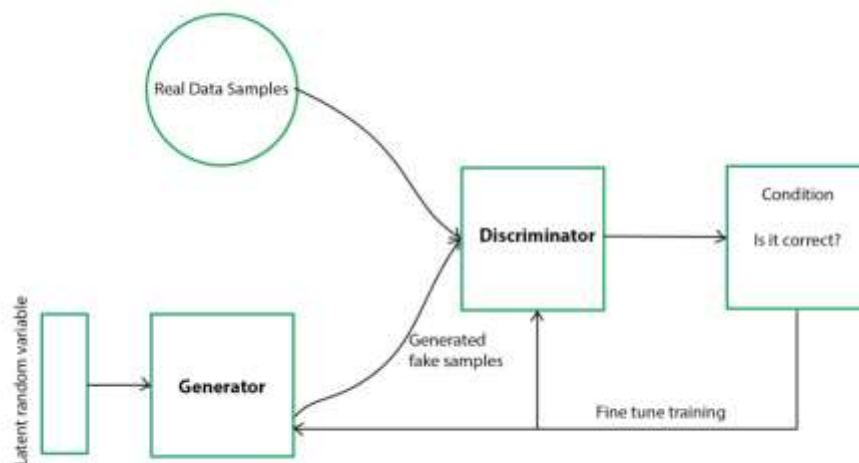


Figure 4.9: Generative Adversarial Networks

Here, the generative model captures the distribution of data and is trained in such a manner that it tries to maximize the probability of the Discriminator in making a mistake. The Discriminator, on the other hand, is based on a model that estimates the probability that the sample that it got is received from the training data and not from the Generator.

The GANs are formulated as a minimax game, where the Discriminator is trying to minimize its reward  $V(\mathbf{D}, \mathbf{G})$  and the Generator is trying to minimize the Discriminator's reward or in other words, maximize its loss. It can be mathematically described by the formula below:

$$\min_G \max_D V(D, G)$$

$$V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$

where,

G = Generator

D = Discriminator

Pdata(x) = distribution of real data

P(z) = distribution of generator

x = sample from Pdata(x)

z = sample from P(z)

D(x) = Discriminator network

G(z) = Generator network

So, basically, training a GAN has two parts:

- **Part 1:** The Discriminator is trained while the Generator is idle. In this phase, the network is only forward propagated and no back-propagation is done. The Discriminator is trained on real data for n epochs, and see if it can correctly predict them as real. Also, in this phase, the Discriminator is also trained on the fake generated data from the Generator and see if it can correctly predict them as fake.
- **Part 2:** The Generator is trained while the Discriminator is idle. After the Discriminator is trained by the generated fake data of the Generator, we can get its predictions and use the results for training the Generator and get better from the previous state to try and fool the Discriminator.

The above method is repeated for a few epochs and then manually check the fake data if it seems genuine. If it seems acceptable, then the training is stopped, otherwise, its allowed to continue for few more epochs.

### Different types of GANs:

GANs are now a very active topic of research and there have been many different types of GAN implementation. Some of the important ones that are actively being used currently are described below:

1. **Vanilla GAN:** This is the simplest type GAN. Here, the Generator and the Discriminator are simple multi-layer perceptrons. In vanilla GAN, the algorithm is really simple, it tries to optimize the mathematical equation using stochastic gradient descent.

2. **Conditional GAN (CGAN):** CGAN can be described as a deep learning method in which some conditional parameters are put into place. In CGAN, an additional parameter 'y' is added to the Generator for generating the corresponding data. Labels are also put into the input to the Discriminator in order for the Discriminator to help distinguish the real data from the fake generated data.
3. **Deep Convolutional GAN (DCGAN):** DCGAN is one of the most popular also the most successful implementation of GAN. It is composed of ConvNets in place of multi-layer perceptrons. The ConvNets are implemented without max pooling, which is in fact replaced by convolutional stride. Also, the layers are not fully connected.
4. **Laplacian Pyramid GAN (LAPGAN):** The Laplacian pyramid is a linear invertible image representation consisting of a set of band-pass images, spaced an octave apart, plus a low-frequency residual. This approach uses multiple numbers of Generator and Discriminator networks and different levels of the Laplacian Pyramid. This approach is mainly used because it produces very high-quality images. The image is down-sampled at first at each layer of the pyramid and then it is again up-scaled at each layer in a backward pass where the image acquires some noise from the Conditional GAN at these layers until it reaches its original size.
5. **Super Resolution GAN (SRGAN):** SRGAN as the name suggests is a way of designing a GAN in which a deep neural network is used along with an adversarial network in order to produce higher resolution images. This type of GAN is particularly useful in optimally up-scaling native low-resolution images to enhance its details minimizing errors while doing so.

#### **Questions:**

1. Define ANN and Neural computing.
2. List some applications of ANNs.
3. What are the design parameters of ANN?
4. Explain the three classifications of ANNs based on their functions. Explain them in brief.
5. Write the differences between conventional computers and ANN.
6. What are the applications of Machine Learning .When it is used.
7. What is deep learning , Explain its uses and application and history.
8. What Are the Applications of a Recurrent Neural Network (RNN)?
9. What Are the Different Layers on CNN?
10. Explain Generative Adversarial Network.

## Unit-5

### 5.1 Image and face recognition

**Image Recognition** is the ability of a computer powered camera to identify and detect objects or features in a digital image or video. It is a method for capturing, processing, examining, and sympathizing images. To identify and detect images, computers use machine vision technology that is powered by an artificial intelligence system. Image recognition is a term for computer technologies that can recognize certain people, animals, objects or other targeted subjects through the use of algorithms and machine learning concepts. The term “image recognition” is connected to “computer vision,” which is an overarching label for the process of training computers to “see” like humans, and “image processing,” which is a catch-all term for computers doing intensive work on image data.

Image recognition is done in many different ways, but many of the top techniques involve the use of convolutional neural networks to filter images through a series of artificial neuron layers. The convolutional neural network was specifically set up for image recognition and similar image processing. Through a combination of techniques such as max pooling, stride configuration and padding, convolutional neural filters work on images to help machine learning programs get better at identifying the subject of the picture.

Image recognition has come a long way, and is now the topic of a lot of controversy and debate in consumer spaces. Social media giant Facebook has begun to use image recognition aggressively, as has tech giant Google in its own digital spaces. There is a lot of discussion about how rapid advances in image recognition will affect privacy and security around the world.

#### **How does image recognition work?**

How do we train a computer to tell one image apart from another image? The process of an image recognition model is no different from the process of machine learning modeling. I list the modeling process for image recognition in Step 1 through 4.

Step 1: Extract pixel features from an image



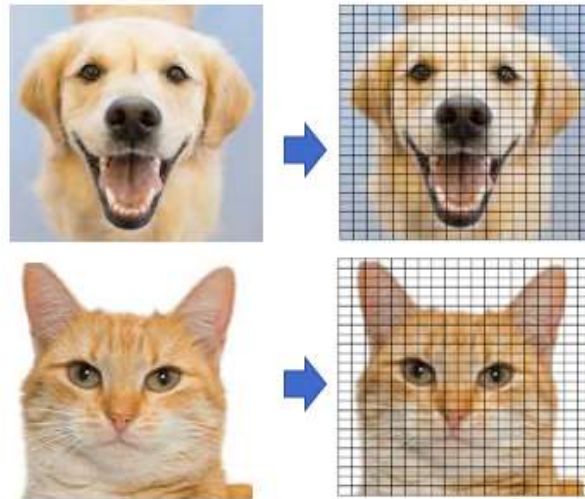


Figure 5.1: Image Recognition

Step 2: Prepare labeled images to train the model

Step 3: Train the model to be able to categorize images

Step 4: Recognize (or predict) a new image to be one of the categories

### Face Recognition

Face recognition is a method of identifying or verifying the identity of an individual using their face. Face recognition systems can be used to identify people in photos, video, or in real-time. Law enforcement may also use mobile devices to identify people during police stops.

But face recognition data can be prone to error, which can implicate people for crimes they haven't committed. Facial recognition software is particularly bad at recognizing African Americans and other ethnic minorities, women, and young people, often misidentifying or failing to identify them, disparately impacting certain groups.

Additionally, face recognition has been used to target people engaging in protected speech. In the near future, face recognition technology will likely become more ubiquitous. It may be used to track individuals' movements out in the world like automated license plate readers track vehicles by plate numbers. Real-time face recognition is already being used in other countries and even at sporting events in the United States.



## How Face Recognition Works

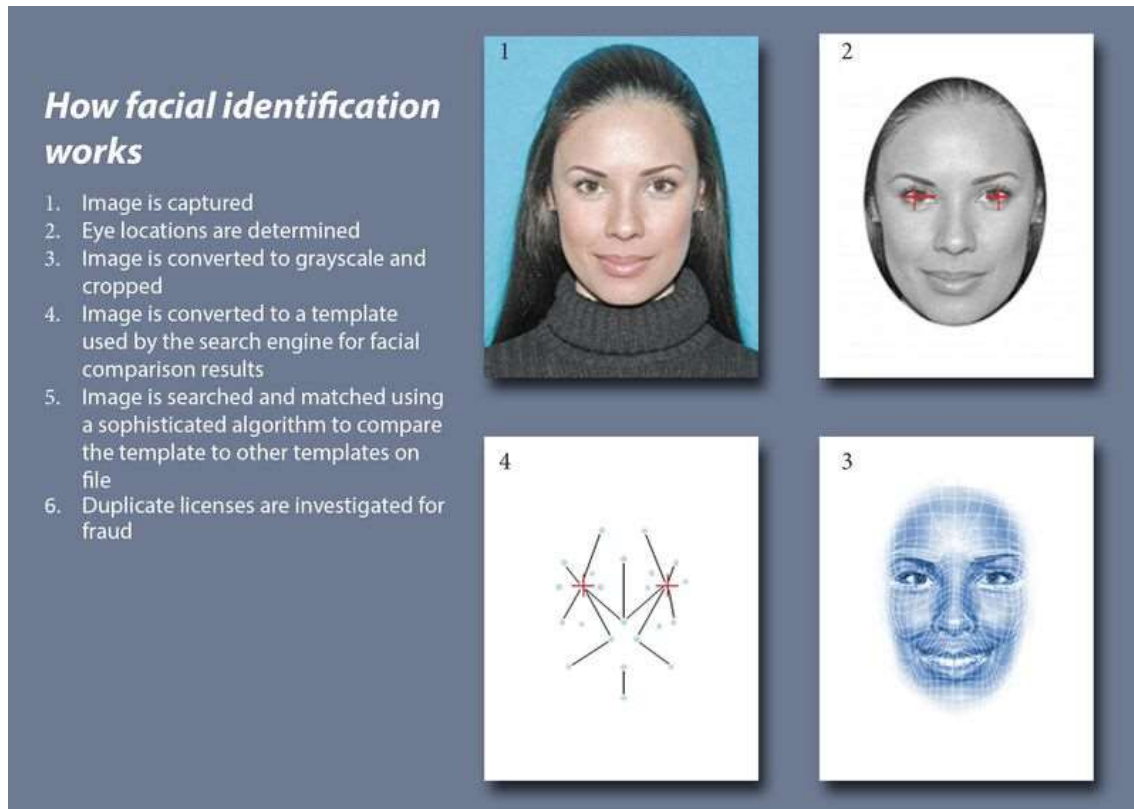


Figure 5.2: Face Recognition

Face recognition systems use computer algorithms to pick out specific, distinctive details about a person's face. These details, such as distance between the eyes or shape of the chin, are then converted into a mathematical representation and compared to data on other faces collected in a face recognition database. The data about a particular face is often called a face template and is distinct from a photograph because it's designed to only include certain details that can be used to distinguish one face from another.

Some face recognition systems, instead of positively identifying an unknown person, are designed to calculate a probability match score between the unknown person and specific face templates stored in the database. These systems will offer up several potential matches, ranked in order of likelihood of correct identification, instead of just returning a single result.

Face recognition systems vary in their ability to identify people under challenging conditions such as poor lighting, low quality image resolution, and suboptimal angle of view (such as in a photograph taken from above looking down on an unknown person).

When it comes to errors, there are two key concepts to understand:

A “false negative” is when the face recognition system fails to match a person’s face to an image that is, in fact, contained in a database. In other words, the system will erroneously return zero results in response to a query.

A “false positive” is when the face recognition system does match a person’s face to an image in a database, but that match is actually incorrect. This is when a police officer submits an image of “Joe,” but the system erroneously tells the officer that the photo is of “Jack.”

## 5.2. Object Recognition

Object recognition is a computer vision technique for identifying objects in images or videos. Object recognition is a key output of deep learning and machine learning algorithms. When humans look at a photograph or watch a video, we can readily spot people, objects, scenes, and visual details. The goal is to teach a computer to do what comes naturally to humans: to gain a level of understanding of what an image contains.

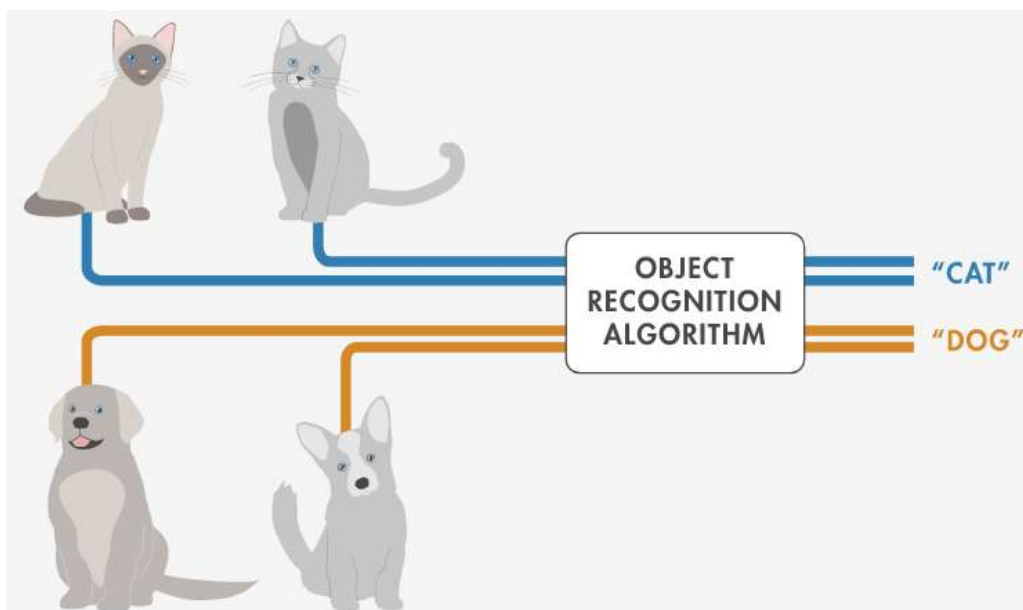


Figure 5.3: Object Recognition

Object recognition is a key technology behind driverless cars, enabling them to recognize a stop sign or to distinguish a pedestrian from a lamppost. It is also useful in a variety of applications such as disease identification in bioimaging, industrial inspection, and robotic vision.

Object detection and object recognition are similar techniques for identifying objects, but they vary in their execution. Object detection is the process of finding instances of objects in images. In the

case of deep learning, object detection is a subset of object recognition, where the object is not only identified but also located in an image. This allows for multiple objects to be identified and located within the same image.

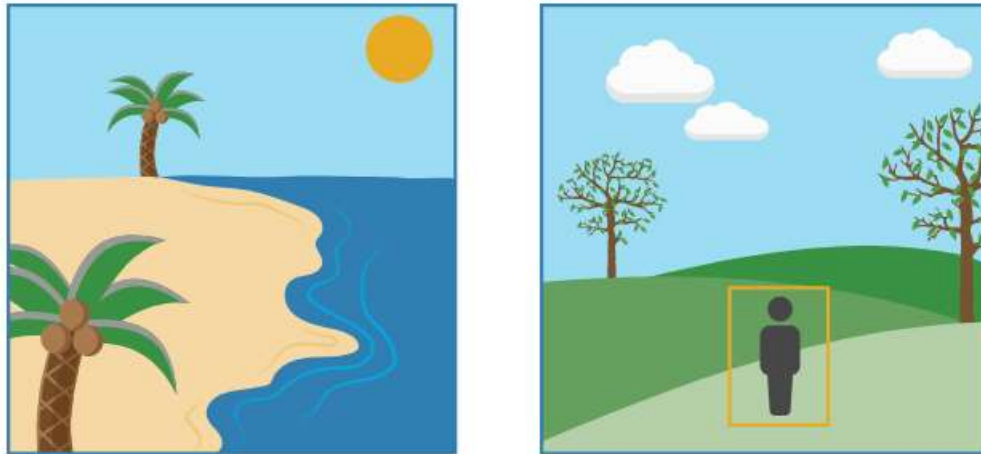


Figure 5.4: Object recognition (left) and object detection (right).

### 5.3. Speech Recognition

Speech recognition, or speech-to-text, is the ability for a machine or program to identify words spoken aloud and convert them into readable text. Rudimentary speech recognition software has a limited vocabulary of words and phrases, and it may only identify these if they are spoken very clearly. More sophisticated software has the ability to accept natural speech, different accents and languages.

Speech recognition incorporates different fields of research in computer science, linguistics and computer engineering. Many modern devices or text-focused programs may have speech recognition functions in them to allow for easier or hands-free use of a device.

It is important to note the terms *speech recognition* and *voice recognition* are sometimes used interchangeably. However, the two terms mean different things. Speech recognition is used to identify words in spoken language. Voice recognition is a biometric technology used to identify a particular individual's voice or for speaker identification.

**How it works**

Speech recognition works using algorithms through acoustic and language modeling. Acoustic modeling represents the relationship between linguistic units of speech and audio signals; language modeling matches sounds with word sequences to help distinguish between words that sound similar.

Often, hidden Markov models are used as well to recognize temporal patterns in speech to improve accuracy within the system. This method will randomly change systems where it is assumed that future states do not depend on past states. Other methods used in speech recognition may include natural language processing (NLP) or N-grams. NLP makes the speech recognition process easier and take less time. N-Grams, on the other hand, are a relatively simple approach to language models. They help create a probability distribution for a sequence.

More advanced speech recognition software will use AI and machine learning. These systems will use grammar, structure, syntax as well as composition of audio and voice signals in order to process speech. Software using machine learning will learn more the more it is used, so it may be easier to learn concepts like accents.

**Applications**

The most frequent applications of speech recognition within the enterprise include the use of speech recognition in mobile devices. For example, individuals can use this functionality in smartphones for call routing, speech-to-text processing, voice dialing and voice search. A smartphone user could use the speech recognition function to respond to a text without having to look down at their phone. Speech recognition on iPhones, for example, is tied to other functions, like the keyboard and Siri. If a user adds a secondary language to their keyboard, they can then use the speech recognition functionality in the secondary language (as long as the secondary language is selected on the keyboard when activating voice recognition. To use other functions like Siri, the user would have to change the language settings.)

Speech recognition can also be found in word processing applications like Microsoft Word, where users can dictate what they want to show up as text.

**Pros and cons**

While convenient, speech recognition technology still has a few issues to work through, as it is continuously developed. The pros of speech recognition software are it is easy to use and readily

available. Speech recognition software is now frequently installed in computers and mobile devices, allowing for easy access.

#### 5.4. Robots

Robotics is a branch of AI, which is composed of Electrical Engineering, Mechanical Engineering, and Computer Science for designing, construction, and application of robots.

##### Aspects of Robotics

- The robots have **mechanical construction**, form, or shape designed to accomplish a particular task.
- They have **electrical components** which power and control the machinery.
- They contain some level of **computer program** that determines what, when and how a robot does something.

##### Difference in Robot System and Other AI Program

Here is the difference between the two –

AI Programs	Robots
They usually operate in computer-stimulated worlds.	They operate in real physical world
The input to an AI program is in symbols and rules.	Inputs to robots is analog signal in the form of speech waveform or images
They need general purpose computers to operate on.	They need special hardware with sensors and effectors.

##### Robot Locomotion

Locomotion is the mechanism that makes a robot capable of moving in its environment. There are various types of locomotions –

- Legged
- Wheeled
- Combination of Legged and Wheeled Locomotion
- Tracked slip/skid

#### Legged Locomotion

- This type of locomotion consumes more power while demonstrating walk, jump, trot, hop, climb up or down, etc.
- It requires more number of motors to accomplish a movement. It is suited for rough as well as smooth terrain where irregular or too smooth surface makes it consume more power for a wheeled locomotion. It is little difficult to implement because of stability issues.
- It comes with the variety of one, two, four, and six legs. If a robot has multiple legs then leg coordination is necessary for locomotion.

The total number of possible **gaits** (a periodic sequence of lift and release events for each of the total legs) a robot can travel depends upon the number of its legs.

If a robot has  $k$  legs, then the number of possible events  $N = (2k-1)!$ .

In case of a two-legged robot ( $k=2$ ), the number of possible events is  $N = (2k-1)! = (2*2-1)! = 3! = 6$ .

Hence there are six possible different events –

- Lifting the Left leg
- Releasing the Left leg
- Lifting the Right leg
- Releasing the Right leg
- Lifting both the legs together
- Releasing both the legs together

In case of  $k=6$  legs, there are 39916800 possible events. Hence the complexity of robots is directly proportional to the number of legs.



### Wheeled Locomotion

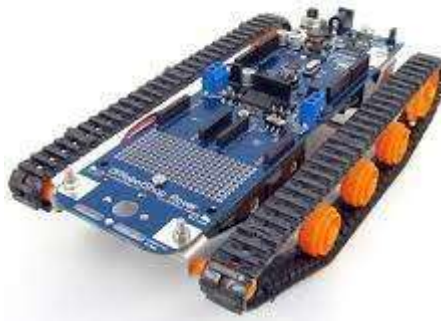
It requires fewer number of motors to accomplish a movement. It is little easy to implement as there are less stability issues in case of more number of wheels. It is power efficient as compared to legged locomotion.

- **Standard wheel** – Rotates around the wheel axle and around the contact
- **Castor wheel** – Rotates around the wheel axle and the offset steering joint.
- **Swedish 45° and Swedish 90° wheels** – Omni-wheel, rotates around the contact point, around the wheel axle, and around the rollers.
- **Ball or spherical wheel** – Omnidirectional wheel, technically difficult to implement.



### Slip/Skid Locomotion

In this type, the vehicles use tracks as in a tank. The robot is steered by moving the tracks with different speeds in the same or opposite direction. It offers stability because of large contact area of track and ground.



### Components of a Robot

Robots are constructed with the following –

- **Power Supply** – The robots are powered by batteries, solar power, hydraulic, or pneumatic power sources.
- **Actuators** – They convert energy into movement.
- **Electric motors (AC/DC)** – They are required for rotational movement.
- **Pneumatic Air Muscles** – They contract almost 40% when air is sucked in them.
- **Muscle Wires** – They contract by 5% when electric current is passed through them.
- **Piezo Motors and Ultrasonic Motors** – Best for industrial robots.
- **Sensors** – They provide knowledge of real time information on the task environment. Robots are equipped with vision sensors to be to compute the depth in the environment. A tactile sensor imitates the mechanical properties of touch receptors of human fingertips.

### 5.5. 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:



### 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.
- 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.

#### Questions:

1. What is the Working of Image Recognition and How it is Used?
2. What is facial recognition - and how sinister is it?
3. What is object recognition in image processing.
4. what is speech recognition in artificial intelligence
5. What's the Difference Between Robotics and Artificial Intelligence?
6. What is robotics?
7. what are applications of artificial intelligence

\*\*\*\*\*The End\*\*\*\*\*