# Paper Title: Machine Learning and Deep Learning with AI

**Radhika Sreedharan**

**Assistant Professor**

**Department of CSE**

**Presidency University, Bangalore**

**radhika.sreedharan@presidencyuniversity.in**

**ABSTRACT**

Machine learning entails compatible means which authorizes computers to learn through practice, example and similarity. Artificial neural networks and genetic algorithms are the most common procedures for machine learning. Machine learning can be used in various applications like sentiment analysis, recommendation systems, voice assistants, business, scientific domain etc. Deep learning can be utilized in robotics control, bio-informatics, autonomous navigation, speech recognition etc. Symbolic reasoning, connections modelled on the brain's neurons, evolutionary algorithms that test variation, Bayesian inference and systems that learn by analogy are the main approaches of AI learning. The main aim of this paper is to mention about role of Machine Learning and Deep Learning in Artificial Intelligence

## I. INTRODUCTION

Learning potentialities can ameliorate the expedition of an intelligent system Machine learning means form the premise for adaptive systems.

Machine learning applications:

a. **Sentiment analysis:** This is an implementation of natural language processing (NLP) Here, the documents words are transformed to feelings such as joyful, regretful and annoyed that are expressed using emoticons constructively. For further analysis or product analysis, five stars or one star are spontaneously connected utilizing sentiment analysis program.
b. **Recommendation systems:** These are systems that build individual acquires feasible. Example: Amazon advices buyers for locating inter-connected books and Netflix considers inter-connected movies of our choice.
c. **Voice assistants:** Products such as Amazon Alexa, Microsoft Cortana, Apple Siri, and Google Assistant are illustrations of voice assistants. They grab lecture instructions and execute jobs. These chat-bots arise from machine learning technologies.
d. **Business:** Forecasting debt of a business organization.
e. **Scientific domain:** Finding new galaxies, recognition of congregation of houses on the basis of type of house/ geographical location, earthquake epicenters' recognition, and recognition of comparable land utilization.

Applications of Deep Learning:

a. **Robotic control:** It is a significant enactment which administers robots utilizing Reinforcement Learning. The method of outlining, feeling, organizing and mastering of the robots can be programmed using these algorithms. The robots can master them independently and take action as intelligent system forced with actual alertness and the capability to envisage and reason like a human being.

   Robots can accomplish specific independent jobs, like vehicle driving, picking and placing an object, traveling in surroundings. The techniques utilized in self-driving cars permits a car to travel in usual congestion not depending on human control.

   A lot of scopes is there in domain known as social robotics, which is categorized as human robot interactions (HCI) and cognitive robotics. Human-robot interactions (HCI) entitles robots to recognize

tasks, feelings, gestures and can travel around surroundings along-with humans. Cognitive robotics entitles robots to master and obtain awareness independently by practice and mastering.

**b. Bio-informatics:** The theoretical field that incorporates molecular biology and genetics with computations statistics is known as bio-informatics. It is dealt with learning the pattern and characteristics in biological data for numerous scientific research which includes biomedicine, genomic sequencing, determination of protein structure, micro-array examination etc. Artificial neural networks and deep learning are utilized to explore, analyze and extract beneficial characteristic in biological data for forecasting and categorizing gene and protein oriented data. Machine learning with computational statistics is helpful to identify genes and nucleotides for identifying a disease on the basis of genes. They are utilized to decide biological sequence and scrutinize patterns like RNA, protein-sequence and DNA.

**c. Linear systems and non-linear dynamics:** System engineers are focused to engineer the surroundings. They get empirical data and consolidate with previous awareness to bring about posterior accomplishment. Example, whenever a sensory sub-system brings about data, it is called experimental awareness. These data are progressed utilizing previous knowledge to bring about output measures to control. This needs modelling of the physical components, physical processes or events. Thus, the engineering products necessitates loop to find the job, outline, investigation, assessment and choice. These demands response. This issue can be split into three sub-problems:

i. Classification
ii. Parameter estimation
iii. State estimation

Example, a speed detector discovers cars that move very quickly and is known as parameter estimation problem. Here, objects/physical methods are reported to locate the sensors' final position. The camera is utilized to obtain the picture of the car and license plate. Discovery is particular demonstration of categorization.

State approximation allots a label and requires characteristics of dynamic systems which repeatedly have random inputs. But it is non-identical to categorization because their data is time ordered. Example, the computation of water quantity is known as state estimation because the data is essentially time-ordered.

The estimations and known area are utilized in finding the posterior awareness.

d. **Autonomous navigation:** It is a demanding job to independently traveling a vehicle or robot in a strange surrounding without human advice. Machine learning plays a significant part to control the speed and navigating for independent mobile robot navigation. The vehicle or robot will be able to navigate effectively by adapting its control and speed perpetually while at the same time learning the adjacent circumstances. Machine learning techniques are likely employed for permitting the succeeding production of robotic applications in commercial, industrial, and military applications. Independent navigation is also advanced to a strange or undetermined surrounding using Unmanned Aerial Vehicles (UAVs) which need high level control processes for learning and adjusting themselves to varying states. Supervised machine learning like neural network along-with reinforcement learning can support every independent navigation.

**e. Speech recognition:** Speech techniques are a standard signal processing scope on the basis of new machine learning standards. Machine learning models can speedily grasp from large datasets for accomplishing facial recognition, speech recognition, speech synthesis etc. Automated speech recognition is a job to convert oral messages into labels as format of text. E.g. the spontaneous subtitles while watching a video, a lot of implementation with voice receptive characteristics. Machine learning techniques with computational intelligence grasp with a lot of data for identifying patterns and recognizing human voice without human interference. Humans talk in informalities, abbreviations and acronyms, it takes substantial investigation of natural language for producing copies. 'Automatic speech recognition' utilizes a lot of machine learning techniques like Hidden Markov Model (HMM), active learning, Bayesian learning and deep learning for providing preferable accomplishment and overpower the fundamental challenges in this part.

## II. CATEGORIES OF LEARNING

Learning is broadly categorized on the basis of nature of learning data and association between the learner and the surrounding:

i. Supervised learning
ii. Unsupervised learning
iii. Semi-supervised learning

**i.** **Supervised learning:** In supervised learning, the machine is trained utilizing data which is well "labeled" which means some data is labeled with an accurate response. It can be contradicted to learning which arises in the attendance of a teacher or a supervisor. Supervised learning algorithm grasps through labeled training data and helpful for forecasting results of unpredicted data. In supervised learning, the model defines the effect of one set of observations, called inputs, on another sets of observations, called outputs. The inputs are assumed to be at the beginning and outputs at the end of the causal chain. The models can incorporate mediating variables between the inputs and outputs.

Supervised learning is utilized in real world applications, like face and speech recognition, products or movie recommendations, and sales forecasting.

In supervised learning, learning data comes with explanation, labels, targets, or desired outcomes and the objective is to find a general rule that maps inputs to outputs. This kind of learning data is called labeled data. The learned rule is utilized for labelling new data with unknown outputs. Supervised learning involves building a machine learning model on the basis of labeled samples.

Example, if we build a system to estimate the price of a plot of land or a house based on various features like size, location, and so on, we first need to create a database and label it. We need to teach the algorithm what characteristics correspond to what prices. On the basis of this data, the algorithm will learn how to calculate the price of real estate making use of values of the input features.

Supervised earning deals with learning a function from available training data. Here, a learning algorithm analyzes the training data and produces a derived function that can be used for mapping new examples. Some of the supervised learning algorithms are:

- Logistic Regression
- Neural networks
- Support Vector Machines
- Naïve Bayes classifiers

Common examples of supervised learning include classifying e-mails into spam and not-spam categories, labeling web pages based on their content and voice recognition.

Supervised learning can be further classified into Regression and Classification.

**a.** **Regression:** Regression trains on and forecasts a continuous-valued response. That means it forecasts a single output value utilizing training data. For example, regression is used to forecast real estate prices. The input variables can be locality, size of a house, etc.
**b.** **Classification:** Classification attempts to find the appropriate class label, such as analyzing positive/negative sentiment, male and female persons, benign and malignant tumors, secure and unsecure loans etc.

**ii.** **Unsupervised learning:** In unsupervised learning, model supervision is not required. Rather, permit the model to perform on its own to find information. It is dealt with the unlabeled data.

Unsupervised learning algorithms permit you to do more complicated processing tasks compared to supervised learning. In unsupervised learning, all the observations are assumed to be caused by hidden variables, that is, the observations are presumed to be at the ending of the causal chain. Unsupervised learning is utilized for detecting anomalies, and outliers, like fraud or defective equipment, or to group customers having behaviors alike for a sales campaign. It is the opposite of supervised learning. There is not labeled data in unsupervised learning. When learning data contains only some warnings lacking explanations or labels, then the algorithm has to find the form of the fundamental data, to find latent patterns, or to find out about data description. This type of learning data is known as unlabeled data. If there are a lot of data points, and we want to categorize them into several groups. We may not know the categorization criteria. So, an unsupervised learning algorithm attempts to categorize the given database into a certain number of groups in an ideal way. Unsupervised learning can be further classified into Clustering and Association

a. **Clustering:** It is dealt with discovering a pattern or structure in a cluster of uncategorized data. The main aim of clustering is to divide the examples into clusters into classes. Every class forecasts feature values for the examples in the class. Every clustering has a forecasting error on the forecasting. The best clustering reduces the error. E. g. an intelligent tutoring system may want to cluster student's learning behavior so that strategies which work for one class member can work for other members.
b. **Association:** Association rules permit to start relationships between data objects inside large databases. This unsupervised technique is with reference to finding stimulating interconnections between variables in large databases. E.g. people who buy a new home may buy a new furniture. Unsupervised learning algorithms are tremendously effective tools to analyze data and to identify patterns and trends. They are utilized for clustering similar input into logical groups. Unsupervised learning algorithms include,
    - K-means
    - Random forests
    - Hierarchical clustering

**Semi-supervised learning:** Semi-supervised learning means some learning samples are labeled and some are nor labeled. It utilizes a large quantity of unlabeled data to train and a small quantity of labeled data to test. Semi-supervised learning is implemented in instances where it is costly to need a fully labeled dataset, whereas it is more practicable to label a small subset.

E.g. it need skilled experts for labelling effective remote sensing images, and lots of field experiments to find oil at a specific location.

## III. APPROACHES TO AI LEARNING

**Symbolic reasoning:** The foremost people, the symbolists, thought that awareness can be achieved by operating on symbols (signs that stand for a certain meaning or event) and extracting rules from them. If the complex system of rules is put together, you can achieve a logic deduction of the outcome you wanted to know, thus the symbolists framed their algorithms for producing rules through data.

*Deduction* enlarges the area of human awareness, while *induction* uplifts the human awareness levels in symbolic reasoning, Induction opens new fields of exploration. Deduction traverses those fields.

**Evolutionary algorithms that test variation:** The evolutionary depends on evolution principles for solving issues. This strategy is on the basis of survival of the fittest (removing any solutions that don't match the desired output). A fitness function finds out the feasibility of every function to solve an issue. The solution method focusses on the best solution on the basis of function outcome utilizing the tree structure. The winner of every evolution level obtains to create the next-level functions. The intention is the next level will get nearer for problem solving but cannot solve it fully, that means another level is needed. This specific group depends on recursion and languages which actively holds up recursion for solving problems. An engrossing outcome of this procedure has been algorithms that enlarge: One generation of algorithms in effect makes the next generation.

**Bayesian inference:** A group of scientists, known as Bayesians, understood that unpredictability was the key feature for keeping an eye on and learning wasn't guaranteed, however, it happened as a continual upgrading of earlier trusts that grew more precise. This approach caused the Bayesians to take on the statistical methods and, in specific, derivations from Bayes' theorem, which assists to compute probabilities under particular conditions.

**Systems that grasp by analogy:** The analogyzers utilize kernel machines for recognizing patterns present in data. We can create a solution of a problem by understanding the pattern of one set of inputs and compare that with pattern of a known output. The objective is for using similarity for finding the best solution to a problem. It's the t that determines type of rationality that utilizing a specific solution performed in a given situation earlier; therefore, utilizing that solution for a similar set of situations should also perform. Recommender systems is one of the most identifiable outputs from this group. Example, when you order a product on Amazon, the recommender system comes up with other, related products which you may like to buy.

The main aim of machine learning is for combining the technologies and procedures embraced by the five tribes for creating a single algorithm which can grasp anything.

## IV. NEURAL NETWORK

A neural network can be defined as a representation of understanding on the basis of human brain. The brain has got a thickly correlated set of nerve cells, or basic information-processing units known as neurons. The human brain integrates almost 10 billion nerve cells (neurons) and 60 trillion connections, synapses, surrounded by them. By making use of multiple neurons at the same time, the brain will be able to carry out its tasks very quickly than speedy computers.

Although every neuron has a very easy form, any array of such elements, comprises an enormous processing power. A neuron consists of a cell body known as **soma**, a lot of fibers known as **dendrites,** and a single long fiber known as **axon.** Dendrites branch into a network all over soma, the axon extends to the dendrites and somas of other neurons.

Signals are spread from one neuron to another using composite electro-chemical reactions. Chemical materials liberated through the synapses give rise to interchange in the electrical potential of the cell body. Whenever the power outreaches its starting point, an electrical pulse, action potential, is sent down from the axon. The pulse scatters and finally reaches synapses, resulting the potential to increase or decrease. But, the engrossing discovery is that a neural network demonstrates flexibility. Neurons exhibit long-term changes in the power of their inter-connection in reaction to the stimulation pattern. Neurons can also form new links along-with remaining neurons. The whole assembly of neurons can move from one place to another. These processes form the learning basis in the brain

Our brain can be thought as very complex. Information can be saved and handled in a neural network at the same time around the whole network, rather than at specific positions. Both data and its processing are global rather than local in case of neural networks.

## V. DEEP LEARNING

The next augmentation in neural networks caused deep learning after backpropagation. Research resumed regardless of AI winter, and neural networks succeeded to deal with technical problems, like *vanishing gradient,* that restricts the neural networks 'dimensions. Developers required larger neural networks for solving some problems which were very large and were unimaginable in the 1980s. Moreover, researchers began to take benefits of the evolution in CPUS and GPUs.

The vanishing gradient is whenever you attempt to send a signal using a neural network and the signal fades to near zero values quickly; it will not get through the activation functions anymore. This takes place as neural networks are chained multiplications. Every below-zero multiplication reduces the values speedily, and activation functions require enough large values so that the signal can pass. If the neuron layers are far from the output, the likelihood will be higher and they'll get locked out of updates as the signals are too small and they

will be stopped by activation functions. As a consequence, your network stops learning as a whole, or it learns very slowly.

New solutions help avoid the problem of vanishing gradient and many other technical problems, and they allow larger deep networks in contradiction to the simpler shallow networks used earlier.

The powerful matrix and vector calculation computing units required for back propagation are GPUs (Graphic Processing Units). These technologies build training neural networks obtainable in less time and approachable to many people. Neural networks can learn from huge amounts of data and take benefits of big data (images, text, transactions, and social media data), creating models that consistently work well, relying on the data flow you feed them.

## VI. DEEP LEARNING IN AI

**Finding even smarter solutions:** Deep learning impacts powers of AI to solve issues in image recognition, machine translation, and speech recognition, that were at first addressed by classic AI and machine learning.

Additionally, it presents new and beneficial solutions:

- Continuous learning using online learning
- Reusable solutions using transfer learning
- More democratization of AI using open-source frameworks
- Simple straightforward solutions using end-to-end learning

**Online learning:** Compared to other machine learning algorithms, neural networks are more adaptable, and they can play a part to train as they perform on making forecasting and categorization. This capacity derives out of optimization algorithms which permit neural networks to learn, that can perform again and again on small samples of examples (known as batch learning) or even on single examples (known as online learning).

Deep learning networks can erect their awareness step by step and be responsive to new knowledge that may appear. For example, a deep learning application on a social media website which can be trained on images of cats. When people post cats' pics, the application identifies them and labels them with a proper label. When people start posting dogs' pics on the social network, the neural network doesn't have to restart training; it can resume by learning dogs' images as well. This ability is especially helpful to manage with Internet data variance. A deep learning network can be open to originality and adjusts its weights to tackle with it.

**Transfer learning:** Adaptability is practical even when a network completes its training, however it should be re-used for goals different from the initial learning. Networks which differentiate objects and accurately categorize them need more time and more of computative ability to learn what to do. Enlarging a network's ability to new types of images which were not part of the earlier learning means transferring the knowledge to this new problem (transfer learning).

For example, we can shift a network which can differentiate between dogs and cats to carry out a task which requires identity cheese and macaroni dishes. We make use of most of network layers as they are and then work on the final, output layers (fine-tuning). The network will apply what it learned in differentiating between dogs and cats to macaroni and cheese. It will work in a better way than a neural network trained only to identify between macaroni and cheese.

**Democratization by using open-source frameworks:** Networks can be accessed by everybody, incorporating approach to tools to create deep learning networks. Scripting abilities are better today as there is a large array of open source deep learning frameworks, such as TensorFlow by Google or PyTorch by Facebook. Those frameworks permit the duplication of the current approaches in deep learning utilizing straight forward commands.

Neural networks require vast quantity of data to work, and data is not approached to all as larger companies hold it. Transfer Learning can reduce insufficient data, but slightly, as certain applications need actual data. As a consequence, the democratization of AI is restricted. In addition, deep learning systems are so complicated that their outcomes are both hard to describe and weak as tricks can mislead the systems. Any neural network can be responsive to adversarial attacks, that are input manipulations formulated to fool the system into giving an incorrect feedback.

**Using end-to-end learning:** Deep learning permits end-to-end learning. It resolves issues in a simple and more uncomplicated way than previous deep learning solution, that may cause considerable influence while solving problem. Consider a problem like AI to identify known faces or drive a car. Utilizing the classical AI approach, the problem to be split into more tractable sub-problems for attaining an allowable outcome in a possible time. Example, if you want to identify faces into a photograph, earlier AI systems arranged the problem into these parts:

- Locate the faces in the photo.
- Trim the faces from the photo.
- Process the trimmed faces to have a pose close to an ID card photo.
- Feed the processed trimmed faces as learning examples to a neural network for image identification.

**REFERENCES**

[1] Artificial intelligence dummies by John Paul Mueller

[2] Artificial intelligence by SaraKutty