International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 05 | May -2020 ISSN: 2582-3930

A VERY BRIEF INTRODUCTION TO MACHINE LEARNING & DEEP LEARNING

Prof. Parinita J Chate, Assistant Professor,
Computer Engineering Department,
Bharati Vidyapeeth's College of Engineering, Lavale, Pune, India

ABSTRACT: The possibility of this research paper is to create attentiveness among upcoming scholars about recent advances in technology, specifically deep learning an area of machine learning which finds applications in big data analytics and artificial intelligence.

Machine learning, especially its subfield of Deep Learning, had many amazing advances in the recent years, and important research papers may lead to breakthroughs in technology that get used by billions of people given the unprecedented availability of data and computing resources, there is widespread renewed interest in applying data-driven machine learning methods to problems for which the development of conventional engineering solutions is challenged by modelling or algorithmic deficiencies.

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

I. INTRODUCTION

Machine learning is a scientific study of algorithms and statistical models that computer systems use to perform a specific task without using explicit instructions. Machine learning algorithms build a mathematical model based on sample data, known as "training data"

Machine learning is closely related to computational statistics, which focuses on making predictions using computers. The study of mathematical optimization delivers methods, theory applications domains to the field of machine learning.

Machine learning algorithms are used in a wide variety of applications, such as email filtering and computer vision, where it is difficult or infeasible to develop a conventional algorithm for effectively performing the task. Today's Artificial Intelligence has far surpassed the hype of block chain and quantum computing. This is due to the fact that huge computing resources are easily available to the common man. The developers

now take advantage of this in creating new Machine Learning models and to retain the existing models for better performance and results.

Machine Learning is a concept which allows the machine to learn from examples and experience, and that too without being explicitly programmed. So instead of writing the code, it feeds data to the generic algorithm, and the algorithm/ machine builds the logic based on the given data.It enables the computers or the machines to make data-driven decisions rather than being explicitly programmed for carrying out a certain task. These programs or algorithms are designed in a way that they learn and improve over time when are exposed to new data.

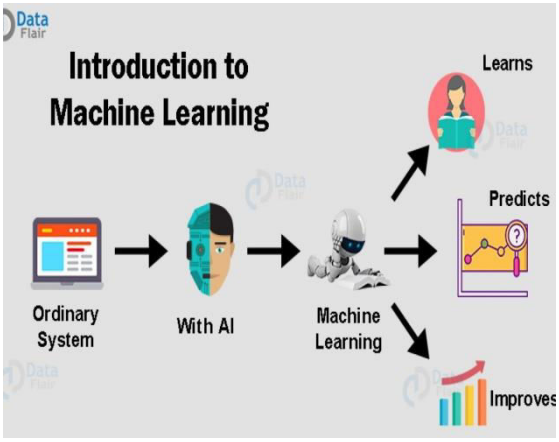


Fig 1 introduction to machine learning

II. HISTORY OF MACHINE LEARNING

Arthur Samuel: He is an American pioneer in the field of computer gaming and artificial intelligence, coined the term “Machine Learning” in 1859 while at IBM. A representative book of the machine learning for pattern classification the interest of machine learning related to pattern recognition continued during the 1970s, as described in the book of Duda and Hart in 1973. In 1981 a report was given on using teaching strategies so that a neural network learns to recognize 40 characters (26 letters, 10 digits and 4 special symbols) from a computer neural. They attempted to approach the problem with various symbolic methods, as well as what were then termed “neural networks”, these were mostly perceptrons and other models that were later found to be reinventions of the generalized linear models of statistics. Probabilistic reasoning was also employed, especially in automated medical diagnosis. He popularized the term "machine learning" in 1959. The Samuel Checkers-playing Program was among the world's first successful self-learning programs, and as such a very early demonstration of the fundamental concept of artificial intelligence (AI). He was also a senior member in the TeX community who devoted much time giving personal attention to the needs of users and wrote an early TeX manual in 1983.

COMPUTER CHECKERS (DRAUGHTS) DEVELOPMENT

Samuel is most known within the AI community for his groundbreaking work in computer checkers in 1959, and seminal research

on machine learning, beginning in 1949.[6] He graduated from MIT and taught at MIT and UIUC from 1946 to 1949.Samuel also designed various mechanisms by which his program could become better. In what he called rote learning, the program remembered every position it had already seen, along with the terminal value of the reward function. This technique effectively extended the search depth at each of these positions. Samuel's later programs reevaluated the reward function based on input from professional games. He also had it play thousands of games against itself as another way of learning. With all of this work, Samuel's program reached a respectable amateur status, and was the first to play any board game at this high a level. He continued to work on checkers until the mid-1970s, at which point his program achieved sufficient skill to challenge a respectable amateur Samuel developed a scoring function based on the position of the board at any given time. This function tried to measure the chance of winning for each side at the given position. It took into account such things as the number of pieces on each side, the number of kings, and the proximity of pieces to being "kinged". The program chose its move based on a minimax strategy, meaning it made the move that optimized the value of this function, assuming that the opponent was trying to optimize the value of the same function from its point of view.

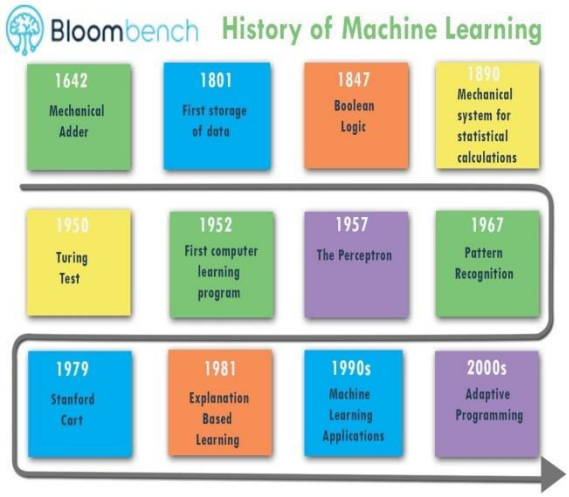


Fig 2 History of Machine Learning - Bloombench - Medium

III. DEEP LEARNING

A new area of machine learning research, which has been introduced with the objective of moving machine learning closer to one of its original goals: Artificial Intelligence.

Deep learning draws its roots from Neocognitron; an Artificial Neuron Network(ANN) introduced by Kunihiko Fukushima in 1980. An ANN is an interconnected network of processing units emulating the network of neurons in the brain. The idea behind ANN was to develop a learning method by modeling the human brain. However, this method lost favor within the machine learning community owing to the fact that it required an impractical amount of time as well as humungous amount of data to train the network parameters for any decent applications. If a machine learning algorithm learns parts of a face like eyes and nose

for face detection task, a deep learning algorithm will learn extra features like the distance between eyes and length of the nose. Here Deep Learning is a major step away from Shallow Learning Algorithms.

A Deep Neural Network (DNN) is defined as an Artificial Neural Network (ANN) with at least one hidden layers of units between the input and output layers. The extra layers give it added levels of abstraction, thus enhancing its modeling capability. The most popular kinds of Deep Learning models, are known as Convolutional Neural Nets (CNN).In recent times, CNNs have also been successfully applied to automatic speech recognition (ASR). Deep Belief Networks and Convolutional Deep Belief Networks are some other popular deep learning architecture in use. There are two disadvantages with DNNs. They are overfitting and computational time. Overfitting is when the DNN learns very specific details on the training data using its hidden layers. As a result, the DNN performs well if the training data is given as input, but poorly when an input data is different. This problem is solved by a method called “dropout” regularization where some units are randomly removed from the hidden layers during training.

The word "deep" in "deep learning" refers to the number of layers through which the data is transformed. More precisely, deep learning systems have a substantial *credit assignment path* (CAP) depth. The CAP is the chain of transformations from input to output. CAPs describe potentially causal connections between input and output. For a feedforward neural network, the depth of the CAPs is that of the network and is the number of hidden layers plus one (as the output layer is also parameterized). For recurrent neural network, in which a signal may propagate through a layer more than once, the CAP depth is potentially unlimited.

The figure below illustrates how categorizing of different images can be achieved using a deep learning model every layer learns a single feature at a time. At the first layer it can learn the different edges; in the second, it could learn slightly more complex features like different parts of a face such as ears, noses and eyes. In the third layer it could learn even more complex features like the distance between eyes or face shapes. The final representations can be used in applications of categorization.

Fig 3 Applications of deep learning

How does deep learning work?

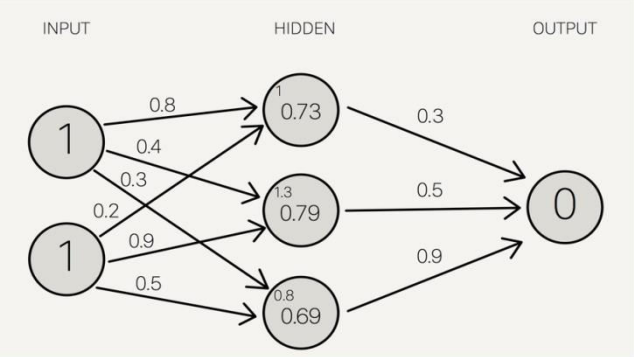
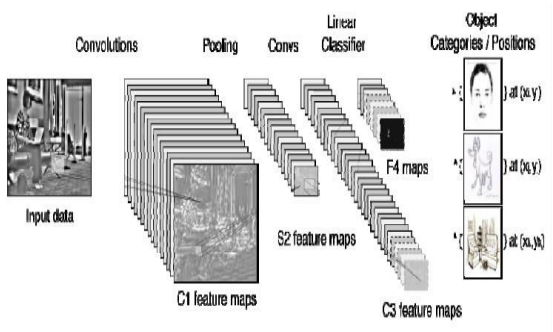


Fig 4 Deep learning how does it work

A deep learning model is designed to continually analyze data with a logic structure similar to how a human would draw conclusions. To achieve this, deep learning applications use a layered structure of algorithms called an **artificial neural network**. The design of an artificial neural network is inspired by the biological neural network of the human brain, leading to a process of learning that's far more capable than that of standard machine learning models. A great example of deep learning is Google's AlphaGo. Google created a computer program with its own neural network that learned to play the abstract board game called Go, which is known for requiring sharp intellect and intuition. By playing against professional Go players, AlphaGo's deep learning model learned how to play at a level never seen before in artificial intelligence, and did without being told when it should make a specific move (as a standard machine learning model would require). It caused quite a stir when AlphaGo defeated multiple world-renowned "masters" of the game—not only could a machine grasp the complex techniques and abstract aspects of the game; it was becoming one of the greatest players of it as well.



So what do machine learning and deep learning mean for customer service?

Many of today's AI applications in customer service utilize machine learning algorithms. They're used to drive self-service, increase agent productivity, and make workflows more reliable. The data fed into those algorithms comes from a constant flux of incoming customer queries, which includes relevant context into the issues those customers, are facing. Aggregating that context into an AI application, in turn, leads to quicker and more accurate predictions. This has made artificial intelligence an exciting prospect for many businesses, with industry leaders speculating that the most practical applications of business-related AI will be for customer service.

IV. DEEP LEARNING IN BIG DATA

Deep Learning and Big Data are two high-focus areas of data science. Deep learning algorithms extract complex data patterns, through a hierarchical learning process by analyzing and learning massive amounts of unsupervised data (Big Data). This makes it an extremely valuable tool for Big Data Analysers.

Big Data has 4 important characteristics, namely, Volume, Variety, Velocity and Veracity. They are Learning algorithms are mainly concerned with issues related to Volume and Variety. Deep Learning algorithms deal with massive amounts of data, i. e. Volume whereas shallow learning algorithms fail to understand complex data patterns which are inevitably present in large data sets.

Deep learning methods are extensively applied to various fields of science and engineering such as speech recognition, image classifications, and learning methods in language processing. Similarly, traditional data processing techniques have several limitations of processing large amount of data. In addition, Big Data analytics requires new and sophisticated algorithms based on machine and deep learning techniques to process data in real-time with high accuracy and efficiency. However, recently, research incorporated various deep learning techniques with hybrid learning and training mechanisms of processing data with high speed.

Semantic Indexing, Data Tagging and Fast Information Retrieval are the main objectives of Deep Learning in Big Data. Consider data that is unstructured and unorganized. Haphazard storage of massive amounts of data cannot be used as a source of knowledge because looking through such data for specific topics of interest and retrieving all relevant and related information would be a tedious task. Using Semantic Indexing and Data Tagging, we identify patterns in the relationships between terms and concepts based on the principle that words used in the same context have similar meanings.

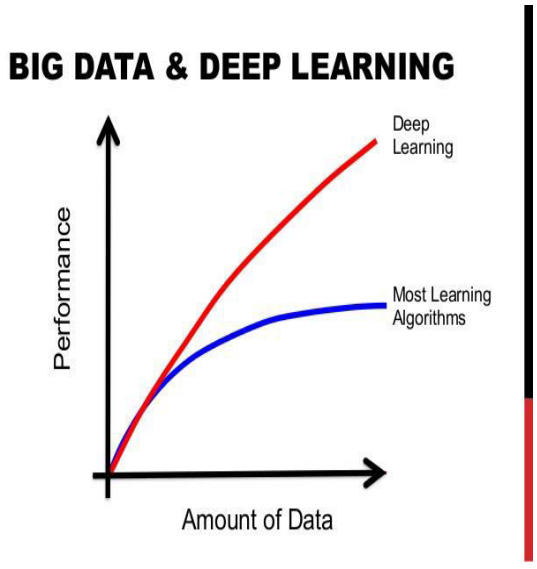
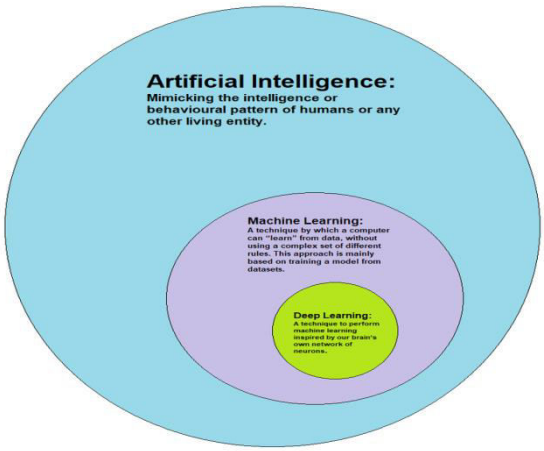


Fig 5 Deep Learning Use Cases - Data Science Pop-up Seattle

V. DEEP LEARNING IN ARTIFICIAL INTELLIGENCE

Artificial Intelligence is the theory and development of computers which are capable of performing tasks which humans can. Deep learning represents the rudimentary level of attempts towards achieving this task. It is utilized in visual perception, speech recognition, game playing, expert systems, decision-making, medicine, aviation and translation between languages.

Fig 6 Artificial intelligence (AI) in deep learning



Artificial intelligence (AI) refers to the simulation of human **intelligence** in machines that are programmed to think like humans and mimic their actions. The term may also be applied to any machine that exhibits traits associated with a human mind such as learning and problem-solving. The ideal characteristic of artificial intelligence is its ability to rationalize and take actions that have the best chance of achieving a specific goal.

Understanding Artificial Intelligence

Artificial intelligence is based on the principle that human intelligence can be defined in a way that a machine can easily mimic it and execute tasks, from the most simple to those that are even more complex. The goals of artificial intelligence include learning, reasoning, and perception.AI is continuously evolving to benefit many different industries. Machines are wired using a cross-disciplinary approach based in mathematics, computer science, linguistics, psychology, and more.

Features

Artificial intelligence refers to the simulation of human intelligence in machines.

The goals of artificial intelligence include learning, reasoning, and perception.

AI is being used across different industries including finance and healthcare

Applications

The applications for artificial intelligence are endless. The technology can be applied to many different sectors and industries. AI is being tested and used in the healthcare industry for dosing drugs and different treatment in patients, and for surgical procedures in the operating room.

Artificial intelligence also has applications in the financial industry, where it is used to detect and flag activity in banking and finance such as unusual debit card usage and large account deposits—all of which help a bank's fraud department. Applications for AI are also being used to help streamline and make trading easier. This is done by making supply, demand, and pricing of securities easier to estimate.

VI. MACHINE LEARNING ADVANTAGES AND DISADVANTAGES:

ADVANTAGES:

Easily identifies trends and patterns

Machine Learning can review large volumes of data and discover specific trends and patterns that would not be apparent to humans. For instance, for an e-commerce website like Amazon, it serves to understand the browsing behaviors and purchase histories of its users to help cater to the right products, deals, and reminders relevant to them. It uses the results to reveal relevant advertisements to them.

No human intervention needed (automation)

With ML, you don't need to babysit your project every step of the way. Since it means giving machines the ability to learn, it lets them make predictions and also improve the algorithms on their own. A common example of this is anti-virus softwares; they learn to filter new threats as they are recognized. ML is also good at recognizing spam.

Continuous Improvement

As **ML algorithms** gain experience, they keep improving in accuracy and efficiency. This lets them make better decisions. Say you need to make a weather forecast model. As the amount of data you have keeps growing, your algorithms learn to make more accurate predictions faster.

Handling multi-dimensional and multi-variety data

Machine Learning algorithms are good at handling data that are multi-dimensional and multi-variety, and they can do this in dynamic or uncertain environments.

DISADVANTAGES:

Data Acquisition

Machine Learning requires massive data sets to train on, and these should be inclusive/unbiased, and of good quality. There can also be times where they must wait for new data to be generated.

Time and Resources

ML needs enough time to let the algorithms learn and develop enough to fulfill their purpose with a

considerable amount of accuracy and relevancy. It also needs massive resources to function. This can mean additional requirements of computer power for you.

Interpretation of Results

Machine Learning is autonomous but highly susceptible to errors. Suppose you train an algorithm with data sets small enough to not be inclusive. You end up with biased predictions coming from a biased training set. This leads to irrelevant advertisements being displayed to customers. In the case of ML, such blunders can set off a chain of errors that can go undetected for long periods of time. And when they do get noticed, it takes quite some time to recognize the source of the issue, and even longer to correct it.

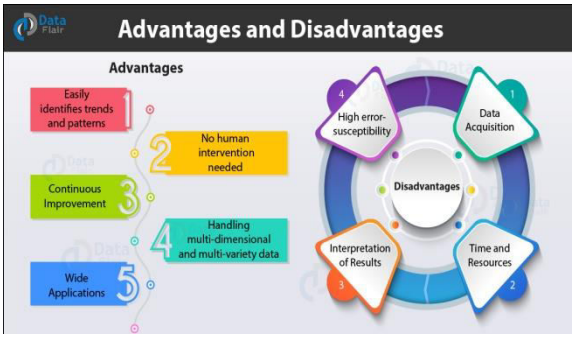


Fig 7 Basic advantage & disadvantage of machine learning

VII. MACHINE LEARNING SKILLS:

Mathematical Notation

Most of the machine learning algorithms is heavily based on mathematics. The level of mathematics that you need to know is probably just a beginner level. What is important is that you should be able to read the notation that mathematicians use in their equations. For

$$f_{AN}(net-\theta)=\begin{cases}\gamma & if\ net-\theta\geq\epsilon\\ net-\theta & if-\epsilon<net-\theta<\epsilon\\ -\gamma & if\ net-\theta\leq-\epsilon\end{cases}$$

$$\max_{\alpha}\left[\sum_{i=1}^{m}\alpha-\frac{1}{2}\sum_{i,j=1}^{m}label^{(i)}\cdot label^{(j)}\cdot a_i\cdot a_j\langle x^{(i)},x^{(j)}\rangle\right]$$

$$f_{AN}(net-\theta)=\left(\frac{e^{\lambda(net-\theta)}-e^{-\lambda(net-\theta)}}{e^{\lambda(net-\theta)}+e^{-\lambda(net-\theta)}}\right)$$

example if you are able to read the notation and comprehend what it means, you are ready for learning machine learning. If not, you may need to brush up your mathematics knowledge.

$$p(c_i|x,y)=\frac{p(x,y|c_i)\,p(c_i)}{p(x,y)}$$

$$\alpha\geq 0, and\ \sum_{i-1}^{m}\alpha_i\cdot label^{(i)}=0$$

Probability Theory

Here is an example to test your current knowledge of probability theory: Classifying with conditional probabilities.

With these definitions, we can define the Bayesian classification rule −

- If P(c1|x, y) > P(c2|x, y) , the class is c1 .
- If P(c1|x, y) < P(c2|x, y) , the class is c2 .

Optimization Problem

Here is an optimization function

$$\max_{\alpha}\left[\sum_{i=1}^{m}\alpha-\frac{1}{2}\sum_{i,j=1}^{m}label^{(i)}\cdot label^{(j)}\cdot a_i\cdot a_j\langle x^{(i)},x^{(j)}\rangle\right]$$

Subject to the following constraints –

Visualization

In many cases, you will need to understand the various types of visualization plots to understand your data distribution and interpret the results of the algorithm's output.

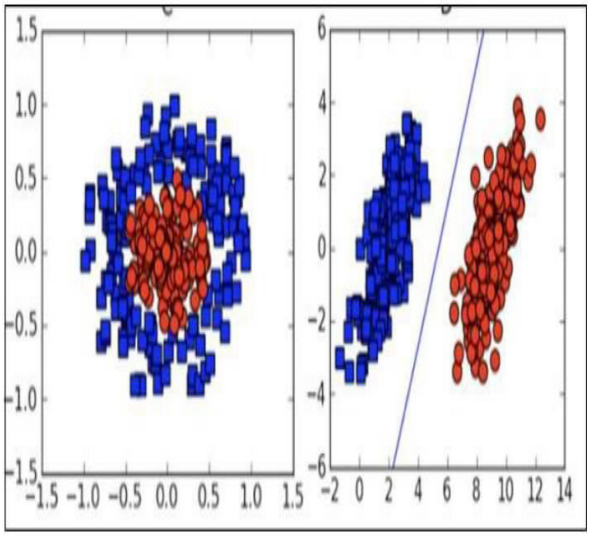


Fig 8 Mathematical Notation of machine learning

VIII. CONCLUSION

Machine Learning is a technique of training machines to perform the activities a human brain can do, albeit bit faster and better than an average human-being, The machines can beat human champions in games such as Chess, AlphaGO, which are considered very complex.

Deep learning techniques have been criticized because there is no way of representing casual relationships (such as between diseases and their symptoms) and the algorithms fail to acquire abstract ideas like "sibling" or "identical to". Not much theory is available for the most of the methods which disadvantageous to beginners. Deep Learning is only a small step towards building machines which have human-like intelligence. Further advancements must be made in order to achieve our ultimate goal. For example, Facebook, Microsoft and Baidu (a Chinese search engine) are buying into this technology and exploring various avenues available.

I consider deep neural networks. Where artificial neural networks have a single hidden layer, deep neural networks have multiple hidden layers. Because of the complexity multiple hidden layers adds to the model, deep neural networks are better at some tasks than simple neural networks.

However, their added complexity m akes them more difficult to train.Then, I discuss convolutional neuron networks. Again, this is a variation of a simple neural network. A benefit to using a convolutional neural network is that it is designed to better handle image and speedch recognition tasks. Instead of hidden layers, convolutional neural networks have a convolutional and pooling layer. It is because of these layers that convolutional neural networks are preferred for image and speech recognition

ACKNOWLEDGEMENT

The authors would like to thank Bharati Vidyapeeth's College of Engineering, Lavale, Pune for supporting and guiding me. Also I am thankful to Principal Sir, Staff who encourage me all the time and willingness to provide assistance whenever it was needed.

REFERENCES

- Aha,D. Kibler,D. (1991): Instance-Based Learning Algorithms. In Machine Learning, vol. 6 - 1 Kluwer Academic Publishers
- Breiman, L., Friedman,J.H., Olshen,R.A. & Stone,C.J. (1984): Classification andRegression Trees, Wadsworth Int. Group, Belmont, California, USA.
- A Deep Learning Tutorial: From Perceptrons To Deep Networks
- An Introduction to Machine Learning Theory And Its Applications: A Visual Tutorial with Examples by Nick Mccrea
- Maryam M Najafabadi, Flavio Villanustre, Taghi M Khoshgoftaar, Naeem Seliya Waldand Edin Muharemagic in"journal of big data".