Cover Page

DOI: http://ijmer.in.doi./2021/10.10.125

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881; IMPACT FACTOR :7.816(2021); IC VALUE:5.16; ISI VALUE:2.286
Peer Reviewed and Refereed Journal: VOLUME:10, ISSUE:10(6), October: 2021
Online Copy of Article Publication Available: www.ijmer.in
Digital certificate of publication: http://ijmer.in/pdf/e-Certificate%20of%20Publication-IJMER.pdf
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: 4th October
Publication Date:30th October 2021

OVERVIEW OF DEEP LEARNING

[1]Roje Spandana Rajeti and [2]Dr. T. Jyothirmayi
[1&2]Assistant Professor
[1]Department of IT and [2]Department of Computer Science and Engineering
[1]UCEV-JNTUK, VZM and [2]GITAM Institute of Technology, VSP (GIT)
Andhra Pradesh, India

Introduction

Deep learning has evolved over time, generating enormous upheaval across sectors and business areaswhichis a subset of mach inelearning that use algorithms for data processing and even constructs abstractions to mimic the thinking process. Deep learning proc esses data using layers of algorithms, interprets human speech, and recognises items visually. Information is transferred through each layer in deep learning, and the output of the previous layer is used as the input for the following layer. The input layers is the first layer in a network, and the output layer is the last. The middle layers are known as hidden layers, and each layer is a simple, uniform algorithm consisting of one type of activation function. Feature extraction, which employs an algorithm to automatically create relevant features of data for learning, training, and comprehension, is another part of deep learning. As a result, deep learning has received a lot of attention and has established itself as a prominent technology in the artificial intelligence community. However, Deep Learning has recently gained prominence as a result of its superior accuracy when trained with large amounts of data. Machine Learning, to put it another way, is a collection of algorithms that analyze data, learn from it, and then use what they've learned to make intelligent decisions.

History/Evolution of Deep Learning

The origins of deep learning may be traced back to 1943, when Warren McCulloch and Walter Pitts developed a computer model based on human brain neural networks. To simulate the mental process, Warren McCulloch and Walter Pitts employed a combination of mathematics and algorithms they called threshold logic. Deep learning has progressed steadily since then, with two important pauses in its development. Henry J. Kelley is credited with developing the fundamentals of a continuous back propagation model in 1960. In 1962, Stuart Dreyfus devised a simplified version based solely on the chain rule. Back propagation was first proposed in the early 1960s, although it was not widely used until 1985.Deep Learning is commonly thought to be a 21st-century creation, yet it has been around since the 1940s, believe it or not. The reason most of us are unaware of Deep Learning advancements/research from the twentieth century is that the methodologies utilized at the time were quite unpopular due to various flaws, and it has undergone several re-brandings since then.

Any new original research in any discipline necessitates a knowledge of the topic's history, evolution, and important breakthroughs that contributed to its popularization. Deep Learning isn't any different. A more comprehensive examination of Deep Learning's history indicates three key rounds of progress:

- During the years 1940–1960, cybernetics was a popular topic.
- During the 1980s and 1990s, connectionism was popular.
- Since 2006, there has been a focus on deep learning.

The first two waves of research were unpopular due to criticisms of their shortcomings, but there is no doubt that they contributed to the advancement of the field to where it is today, and some of the algorithms developed during those times are still used widely in various machine learning and deep learning models today. Let's take a closer look at the three waves to have a better understanding.

Cybernetics

Is the forerunner of modern Deep Learning, and is based on the concept of biological learning, or how human brains learn? The purpose of advances in cybernetics was to recreate the workings of a human/animal brain in a simplified computer model, which would aid in the development of systems that would start learning like real brains and produce conclusions given some input. Until date, research in this mindset has been conducted separately under the umbrella of Computational Neuroscience.

The McCulloch-Pitts Neuron was the catalyst for cybernetics to take off. It was a stab at emulating a biological neuron. It was based on a linear model that took multiple inputs [X1, X2,.... Xn], had some weights [W1, W2,... Wn] for each input, and produced the output f(x,w) = X1W1 + X2W2 +.... + XnWn. Based on the inputs and weights, this model could only produce True/False and the weights had to be adjusted manually.

Cover Page

DOI: http://ijmer.in.doi./2021/10.10.125

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881; IMPACT FACTOR :7.816(2021); IC VALUE:5.16; ISI VALUE:2.286
Peer Reviewed and Refereed Journal: VOLUME:10, ISSUE:10(6), October: 2021
Online Copy of Article Publication Available: www.ijmer.in
Digital certificate of publication: http://ijmer.in/pdf/e-Certificate%20of%20Publication-IJMER.pdf
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: 4th October
Publication Date:30th October 2021

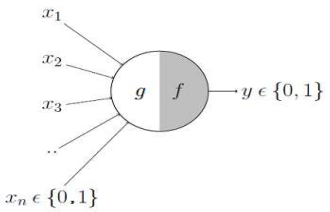

Fig: McCulloch-Pitts Model

Later in the 1950s, an American psychologist named Frank Rosenblatt created Perceptron, which learned the weights automatically. Perceptron was created as an electrical machine rather than a programme or software. For image identification, Frank created the Perceptron, which consisted of photocells (receivers) coupled to numerous neurons that classified the inputs collected by the photocells. Perceptron was an amazing gadget at the time, but it made huge claims that could not be fulfilled.

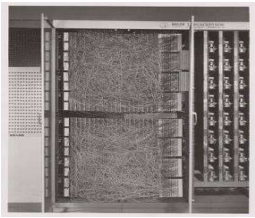

Fig: Perceptron Model

ADALINE - Adaptive linear element, invented by Bernard Widrow about the same time as Perceptron, may also adjust to the weights depending on the weighted sum of the inputs during the learning phase.

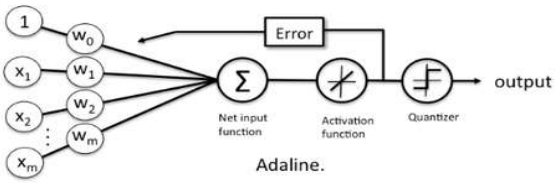

Fig: ADALINE Model

ADALINE's learning function is similar to the stochastic gradient descent employed in today's Linear Regression. These linear models had a number of flaws, and critics who pointed out these flaws caused a significant drop in their popularity, halting development for a time. One significant drawback was that these linear models couldn't be used to train for XOR functions. Because these models were based on neuro-scientific research, the decline in their popularity prompted the development of new models that were not based on neuro-scientific research.

Connectionism Or Parallel Distributed Processing:

This was very popular in the 1980s. Cognitive sciences were the inspiration for this method. In comparison to the many symbolic reasoning techniques known as Classicists that scientists were researching in the 1980s, Connectionism showed promise. Even while Symbolic Reasoning fits well in a more abstract picture of the brain, it is difficult to apply explicitly using standard programming models. As a result, practical connectionists saw their work as a way to use Neural Nets to achieve the same effect as Symbolic Reasoning. However, Radical Connectionists simply dismissed the idea of Symbolic Reasoning, claiming that it couldn't explain various complex elements of our brain and was an inaccurate perception of the human brain in the first place.

During this wave, the concept of Artificial Neural Networks (ANNs) was introduced. The primary goal of ANNs was to create a network of individual units that could be programmed to perform intelligent tasks. The concept of hidden layers was originally proposed at this period. Parallel signal processing was dispersed along several branches of the network thanks to a network of artificial neurons connected to each other. Weights were added to the connections between the "neuron" modules to adjust the strength of the effect one neuron has on another. This technique was thought to be quite comparable to what happens inside human nervous system, which raised questions among academics about the models' usefulness.

Cover Page

DOI: http://ijmer.in.doi./2021/10.10.125

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881; IMPACT FACTOR :7.816(2021); IC VALUE:5.16; ISI VALUE:2.286
Peer Reviewed and Refereed Journal: VOLUME:10, ISSUE:10(6), October: 2021
Online Copy of Article Publication Available: www.ijmer.in
Digital certificate of publication: http://ijmer.in/pdf/e-Certificate%20of%20Publication-IJMER.pdf
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: 4th October
Publication Date:30th October 2021

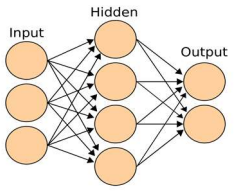

Fig: ANN Model

LSTM, distributed representation and processing, and back-propagation to train deep neural nets were invented during this wave of Connectionism, and they remain fundamental components of several advanced deep learning applications to this day. However, by the mid-1990s, AI-based firms were making unrealistic claims and, because to a lack of computational resources, could never provide that degree of sophistication from these models. Investors backed off, causing a drop in the second wave of deep learning. The second wave did not die, but it did dwindle. Although research was conducted at numerous labs, applications were scarce until the early 2000s.

Deep Learning

After two dips, the third wave finally broke through in 2006. Deep Belief Networks were trained using Greedy Layer-wise Training by Geoffrey Hinton. DBNs are, in their most basic form, a collection of many hidden layers, each of which contains a set of latent variables. Layers have connections, but not between the variables within each layer. Restricted Boltzmann machines are another name for a very simple DBN implementation.

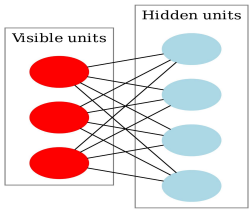

Fig: Restricted Boltzmann Machine

Other researchers exploited Geoffrey Hinton's innovations to train other types of Deep Networks. This allowed academics all over the world to train ever-deeper neural networks, giving rise to the phrase **"Deep Learning."** While it may appear that Geoffrey Hinton was the catalyst for the development of deep learning, the increased processing power and availability of enormous datasets cannot be overlooked. When the same algorithms established during Connectionism were trained on larger and larger datasets, they began to produce better results. The difference between then and today is that, as more people use internet services, we have a lot more data and far better computational capabilities to deal with it, resulting in improved model accuracy.

What is Deep Learning?

Deep learning is an application of machine learning that learns to represent the world as a nested hierarchy of concepts, with each concept defined in relation to simpler concepts and more abstract representations computed in terms of less abstract ones. It achieves great power and flexibility by learning to represent as a nested hierarchy of concepts. Deep learning is a machine learning subfield that works with algorithms that are inspired by the structure and function of the brain. Machine learning is a subset of artificial intelligence, while deep learning is a subset of machine learning (AI).

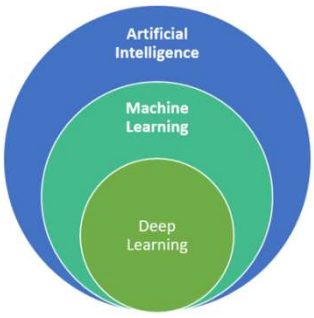

Fig: Relationship b/w AI, ML and DL

Deep learning algorithms can self-learn hidden patterns inside data to create predictions with accelerated processing power and big data sets. Deep learning is a type of machine learning that is trained on massive amounts of data and uses a large number of

Cover Page

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881; IMPACT FACTOR :7.816(2021); IC VALUE:5.16; ISI VALUE:2.286
Peer Reviewed and Refereed Journal: VOLUME:10, ISSUE:10(6), October: 2021
Online Copy of Article Publication Available: www.ijmer.in
Digital certificate of publication: http://ijmer.in/pdf/e-Certificate%20of%20Publication-IJMER.pdf
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: 4th October
Publication Date:30th October 2021

DOI: http://ijmer.in.doi./2021/10.10.125

compute units to make predictions. Deep learning is a type of machine learning in which a computer algorithm learns to execute categorization tasks directly on complicated input such as photos, text, or voice. These algorithms can achieve state-of-the-art (SOTA) accuracy, and in some cases, even outperform humans. They are trained using a vast set of labelled data and multilayer neural network topologies.

Deep Learning is a key component of technologies such as virtual assistants, facial recognition, and self-driving cars. Deep learning is based on the training of data and the learning of experiences. The learning technique is known as 'Deep', since the neural networks rapidly uncover new levels of data with each passing minute. Each time data is trained, the goal is to improve performance. Because it is widely embraced by data professionals, its training performance and deep learning skills have improved dramatically as the depth of the data has increased.

Importance of Deep Learning

Deep learning is capable of performing sophisticated operations on both organized and unstructured data. To discover patterns, deep learning takes in massive amounts of data and analyses it to extract features from an object while maintaining the model's performance. When dealing with unstructured data, deep learning's capacity to process vast amounts of features makes it particularly strong. Deep learning techniques, on the other hand, may be overkill for simpler tasks because they require access to a large amount of data to be effective. ImageNet, for example, is a prominent benchmark for training deep learning models for comprehensive image recognition, with over 14 million images available. It's very easy for a deep learning model to become over-fitted and fail to generalize well to new data if the data is too simple or incomplete. As a result, for most practical business tasks like assessing customer churn, detecting fraudulent transactions, and other scenarios with smaller datasets and fewer characteristics, deep learning models are less effective than other techniques (such as boosted decision trees or linear models). Deep learning can work for smaller, structured datasets in some circumstances, such as multiclass classification.

Advantages of Deep Learning

- The ability to create new features from the limited training data sets.
- The ability to work with unsupervised learning approaches aids in the generation of actionable and consistent task outcomes.
- It decreases the time necessary for feature engineering, which is one of the chores that takes up a lot of time while practicing machine learning;
- Its architecture has become adaptable to change and can work on a variety of problems as a result of ongoing training.
- Best problem-solving performance in class.
- Eliminates unnecessary costs by reducing the requirement for feature engineering.
- Easily detects problems that are otherwise difficult to identify.

Disadvantages of Deep Learning

- As the quantity of datasets grows, the cost of computer training rises dramatically.
- Because the entire training process is dependent on the constant flow of data, there is less room for improvement in the training process.
- There is a lack of openness in the defect revision process. There are no intermediary steps to present justifications for a specific flaw. In order to fix the problem, the entire algorithm is rewritten.
- Expensive resources, high-speed processing units, and strong GPUs are required for data set training.

Challenges in Deep Learning

Though deep learning approaches have exploded in popularity in the last decade or so, the concept has been present since Frank Rosenblatt built the perceptron on an IBM® 704 system in the mid-1950s. It was a two-layer electrical gadget that could recognize shapes and reason about them. The increase in computing power and high-performance graphical processing units (GPUs) in recent years, combined with the large increase in the wealth of data these models have at their disposal for learning, as well as interest and funding from the community for continued research, have all contributed to advancements in this field. Though deep learning has exploded in popularity in recent years, it has its own set of some Limitations and challenges that the community is attempting to overcome:

- Deep learning necessitates a lot of data. In addition, more powerful and accurate models will have more parameters, which will necessitate more data.

Cover Page

DOI: http://ijmer.in.doi./2021/10.10.125

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881; IMPACT FACTOR :7.816(2021); IC VALUE:5.16; ISI VALUE:2.286
Peer Reviewed and Refereed Journal: VOLUME:10, ISSUE:10(6), October: 2021
Online Copy of Article Publication Available: www.ijmer.in
Digital certificate of publication: http://ijmer.in/pdf/e-Certificate%20of%20Publication-IJMER.pdf
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: 4th October
Publication Date:30th October 2021

- Deep learning models become inflexible and incapable of multitasking once they have been trained. They are capable of providing effective and precise answers, but only to a single problem. Even resolving a similar issue would necessitate system retraining.
- Any application that involves thinking, such as programming or following the scientific method, requires long-term planning and algorithms, such as data manipulation, that are utterly beyond the capabilities of existing deep learning approaches, even when dealing with massive amounts of data.

Applications of Deep Learning

Deep learning is commonly used to predict rain, earthquakes, and tsunamis in the weather. It aids in the implementation of appropriate safeguards. Machines can understand speech and produce the required output. It allows machines to distinguish people and things in photographs that are supplied to them. Advertisers can also use deep learning models to execute real-time bidding and targeted display advertising. Apart deep learning has a wide range of applications including health care, finance, image recognition, Autonomous vehicles, e-commerce, Personal assistant, Automatic Text Generation, Aerospace and military, Customer experience (CX), Adding color to photos and videos, Computer vision tasks, Industrial automation.

Archietecture of Deep Learning

Deep learning algorithms use supervised and unsupervised learning algorithms to train outputs based on the inputs provided. In the figure below, the circles indicate linked neurons. Input, Hidden, and Output Layers are three separate hierarchies of layers that the neurons are divided into. The input layer, which is the first neuron layer, receives the data and passes it on to the first hidden layer. The computations are done on the received data by the hidden layers. The most difficult aspect of creating neural networks is deciding on the number of neurons and hidden layers. Finally, the output layer generates the output necessary.

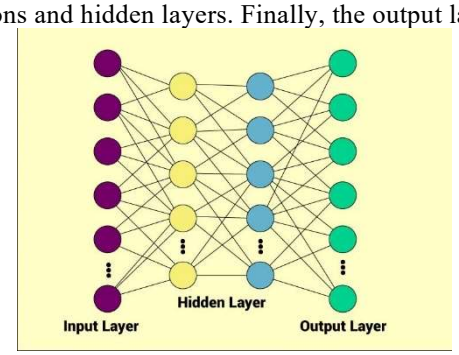

Fig: Working network of Deep Learning

Weights are used to indicate the importance of input values in every link between neurons. An activation function is used to normalize the outputs. Two crucial measures are taken into account when training the network. The first is to gather a great amount of data, and the second is to have a lot of computing power. The term "deep learning" refers to the amount of hidden layers used by the model to train the data set.

Work flow of deep learning

To achieve the proper solution, we must first identify the actual problem, which must be comprehended. The practicality of Deep Learning should also be examined (whether it should fit Deep Learning or not). Second, we must determine the pertinent facts that must match to the actual situation and be prepared properly. Third, select a suitable Deep Learning Algorithm. Fourth, during training the dataset, Algorithm should be employed. Fifth, the dataset should be subjected to final testing.

Cover Page

DOI: http://ijmer.in.doi./2021/10.10.125

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881; IMPACT FACTOR :7.816(2021); IC VALUE:5.16; ISI VALUE:2.286
Peer Reviewed and Refereed Journal: VOLUME:10, ISSUE:10(6), October: 2021
Online Copy of Article Publication Available: www.ijmer.in
Digital certificate of publication: http://ijmer.in/pdf/e-Certificate%20of%20Publication-IJMER.pdf
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: 4th October
Publication Date:30th October 2021

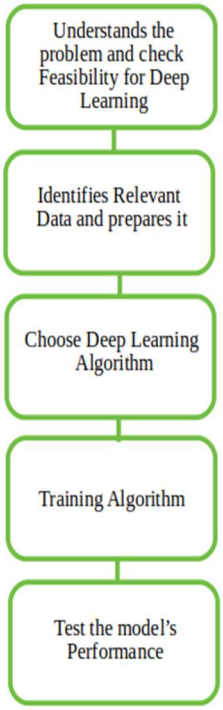

Fig : Work flow of Deep Learning

Conclusion

Although Deep Learning has progressed significantly from its inception in the 1940s, it is critical to recall where it all began and how it has evolved through time because it still has a long way to go. Studying the early building blocks will aid in the development of unique Deep Learning applications in the future. Deep Learning now outperforms in terms of performance and is widely employed for a range of tasks. However, this have not happened in a matter of years but took decades.

References

1. https://www.simplilearn.com/tutorials/deep-learning-tutorial/introduction-to-deep-learning
2. https://developer.ibm.com/articles/an-introduction-to-deep-learning/
3. https://www.analyticssteps.com/blogs/deep-learning-overview-practical-examples-popular-algorithms
4. https://searchenterpriseai.techtarget.com/definition/deep-learning-deep-neural-network
5. https://towardsdatascience.com/the-deep-history-of-deep-learning-3bebeb810fb2
6. https://www.dataversity.net/brief-history-deep-learning/
7. https://www.analyticsinsight.net/the-history-evolution-and-growth-of-deep-learning/
8. https://builtin.com/artificial-intelligence/deep-learning-history

Books

- Deep Learning by Ian Goodfellow
- Deep Learning Revolution by Terry Sejnowski
- Perceptrons by Marvin Minsky and Seymour Papert

Articles

- Solving XOR with a single Perceptron - medium.com
- Deep belief networks - www.scholarpedia.org
- McCulloch-Pitts Neuron — Mankind's First Mathematical Model Of A Biological Neuron - towardsdatascience.com