DEEP LEARNING: EVOLUTION

Suraj Kumar

University institute of Sciences, Chandigarh University, Mohali, Punjab, India

---

ABSTRACT: Deep learning is an old concept, but only a few years ago, neither the term "Deep learning" nor the approach was popular, then the field was reignited by documents such as ImageNet's 2012 deep net model Krizhevsky, Sutskever & Hinton. Deep learning is a growing area of Machine learning. It made of multiple layers of artificial neural networks hidden within. The Deep learning methodology applies nonlinear transformations and high-level model abstractions in large databases. Recent advances in Deep learning have already made important contributions to the field of AI. This paper presents a detailed evolution of Deep learning. This paper describes how and what Deep learning algorithms have been used in major applications. In addition, in common applications, the advantages of the Deep learning methodology and its hierarchy in layers and nonlinear operations are presented and compared to more conventional algorithms.

Keywords: Machine Learning, Deep Learning, Supervised Learning, Unsupervised Learning, Neural Network

INTRODUCTION: Deep learning is a part of Machine learning and AI is been there since a long time. But Deep learning came into light in March 2016, when AlphaGo, an AI bot designed by Google won 4 out of 5 matches against Human Grand master Lee Sedol in Chinese abstract strategy board game Go. Since, then Deep learning training and learning got recognition all around for Humanizing machines.

Fig. 1: Lee Sedol vs AlphaGo: Final Match

The 2018 ACM A.M. Turing Award, was awarded to three of Deep learning's most prestigious architects, Facebook's Yann LeCun, Google's Geoffrey Hinton, and University of Montreal's Yoshua Bengio. Also known as Godfathers of AI and Fathers of Deep learning Revolution. These three along with many others over the past decade, developed the algorithms, systems, and techniques responsible for the aggressive advancements of AI-powered products and services.

Fig. 2: Fathers of deep learning revolution

Most of the advanced potential found nowadays in AI platforms is due to the massive growth of Machine learning and Deep learning technologies. Even though most of the experiments related to Deep learning are in

infancy but researchers predicted that Deep learning will provide a terrifying swiftness in adaption and growth of AI technologies. As previously written, Deep learning is a machine-leaning branch that deploys data processing algorithms and mimics the complex thinking process of the human mind. Deep learning uses algorithm layers to process the data, in understanding human speech and recognizing objects visually. In Deep learning, data passes through many layers and each layer's output operates as the next layer's input, where the first layer is often called the Input Layer, the hidden layer is known as the mid Deep Learning layer and the last layer is called the Output Layer. Each of these layers consists of a single kind of activation function, a simple and consistent algorithm. An alternative characteristic of Deep learning is Feature Extraction, in which an algorithm is used in order to build meaningful data aspects for learning, training, and understanding automatically

2. BACKGROUND: Before discussing further about Deep learning, let us get a grasp of some related terms for better understanding: - – Machine learning Vs AI – AI, what is it? "AI refers to a computer program which is capable of replicating Human thinking process in machines." When AI was introduced, researchers were trying to imitate human intelligence for a certain task. Like playing a game, Researchers introduced a large set of rules that the computer must follow. In AI computer have a specific list of actions that it can execute to perform a task and made decisions on based of predefined set of rules.

– Machine learning, what is it? "Machine learning refers to a machine's ability of learning from a massive data set instead of using coded decision-making rules." Machine learning gives self-learning ability to a computer. Machine learning allows a computer to learn itself without or minimum human interference. This sort of learning takes advantage of modern computers' monstrous processing capabilities, which can very conveniently process massive data sets.

2.1 Supervised Learning Versus Unsupervised Learning

– Supervised Learning, what is it? "Supervised Learning use data which is 'Labeled.' It means some of the data from training set is already tagged with the right answer." A supervised learning algorithm learns from labeled data and enables you to predict unpredicted data outcomes.

– Unsupervised Learning, what is it? "In Unsupervised Learning there's no need to supervise your data science model. Instead, you need to allow your model to function freely and discover information. It deals mainly with unlabeled data." Unsupervised Learning allows you to perform more complex data processing tasks when compared to Supervised Learning. Nonetheless in comparison to other natural learning, Deep learning, and reinforcement learning techniques, unsupervised learning can be unpredictable.

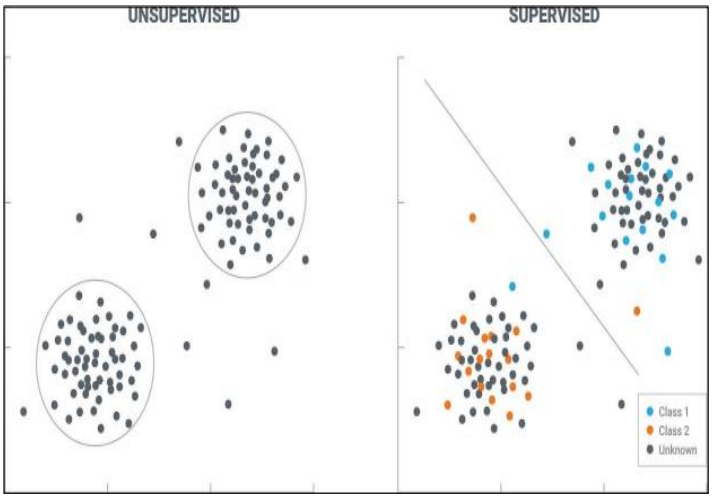

Fig 3: Supervised and Unsupervised Learning graph

3. How Deep learning Works? : AI brain has neurons to process data just like human brain. These neurons are reliant in nature. These neurons are divided into 3-layer types: -

- Input Layer
- Hidden Layers
- Output Layer

– Input data is acquired using the Input Layer. Data is transferred to the initial hidden layer by the input layer.

– Then the Hidden layers perform mathematical calculation on the input data. One of the obstacles in designing neural networks is to decide the number of hidden layer and the number of neurons required for every layer. – In Deep learning, the word 'Deep' means having multiple hidden layers.

– The Output Layer returns the result of the processed data

Every link between neurons has a certain weight. The weight tells the model how significant the value of an input is. The weights are set up irregularly.

There is an Activation Function for every neuron. In the absence of mathematical reasoning, activation functions

are difficult to comprehend. Standardizing outputs is one purpose of the activation functions. When every individual layer of the Neural Network has gone through a group of input data, it outputs the result from the Output Layer as output data.

4. Progression of Deep learning: The origin of Deep learning is dated to 1943, when Warren McCulloh and Walter Pitts designed a model inspired by the Neural Network of the Human Brain.

They combined the algorithms and mathematics they named Threshold Logic to imitate a human brain with the thought process. Since then, with two significant breakthroughs in its growth, Deep learning has evolved tremendously.

– In 1960, Henry J. Kelly developed the basics of continuous Back Propagation.

– A fundamental version of Back Propagation was designed by Stuart Dreyfus, in 1962, which was built on Chain Rule. As mentioned above method Back Propagation was present in 1960 but it was clumsy and not as efficient as its today, and it was not useful till 1985.

– In 1965, Alexey Grigoryevich Ivakhnenko and Valentin Grigor'evich Lapa used polynomial activation function in their model's layer. Then they were analyzed statistically. Then from every individual layer the best statistical options are transferred to the next layer for further processing.

– In 1970's the first AI winter taken place, because of the promises that can't be fulfilled. The lack of funding was hitting AI and Deep learning and that results in limited research options for AI and Deep learning researchers. Luckily, there were some researchers who continued the research, in the absence of funding.

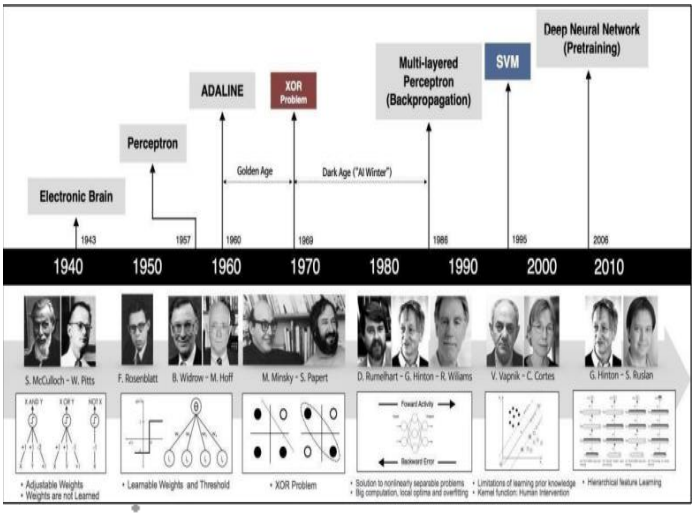

Fig 4: Evolution of Deep Learning

– In 1970, Back Propagation evolved tremendously, using errors during the training of Deep learning models. This occurred when his thesis which included a FORTRAN code for back propagation, which had been written by Seppo Linnainmaa. Unfortunately, this notion did not come into use until 1985. It was in 1985 when D. Rumelhart, R. Williams, and G. Hinton showed that back propagation in a neural network can provide many interesting representations of distribution.

– Kunihiko Fukushima first used Convolutional Neural Networks, or CNN, in 1979. The concept was to design multiple pooling and convolutional layers of a neural network. A neural network making use of a hierarchical, having multiple layers design was introduced by Fukushima. This design has enabled computers to learn how visual patterns can be recognized. This design was named "Necognitron" by Fukushima. The networks of Fukushima were like modern neural networks, but trained on many layers with a reinforced strategy of recurring activation. The networks of Fukushima made it possible to manually adjust the features by increasing the weight of certain links.

– In 1989, at Bell Labs, Yann LeCunn came up with the first demonstration of back propagation in practical world. By combining Convolutional Neural Networks with back propagation, he designed a model to read handwritten numbers. Ultimately, this scheme is used for reading handwritten numbers.

– The second winter of AI took place in the Deep learning and Neural Networks studies of the 1985-1990s. This was the result of some overly optimistic researchers making exaggerated promises to investors, resulting in breaking expectations and angering investors. The situation was so

critical that the status of Pseudoscience reached AI. Luckily, some of the researchers carried on their AI and Deep learning research and much significant progress was made.

– In 1995, D. Cortes and V. Vapnik introduced a system called the Support Vector Machine for recognizing and mapping similar data. – Long Short-Term Memory (LSTM) for RNN was designed by Sepp Hochreiter and Juergen Schmidhuber in 1997.

– In 1999, GPUs were developed and data processing computers became faster. In Deep learning, this resulted in important evolutionary growth. These developments have caused a spike in computational speed by 1000 times over a decade. Neural networks initiated to compete against support vector machines due to the above modifications. Neural networks can be slow in comparison with support vector machines, but neural networks give improved results with the same data. Neural Networks also possess another merit over the support vector machine, which is as more training data is provided, it can always improve.

– In 2000, the upper layers were found not to be learning characteristics or lessons created in lower layers, as learning signals are unable to reach these layers. This is known as The Problem of the Vanishing Gradient. In all the neural networks available, the given problem doesn't exist, only in those that use of gradient-based learning methods. Certain activation functions are the cause of this issue. Their inputs were precipitated by some activation functions, which in turn messily reduced the output range. The effect is that a huge input area has been mapped over a small range. An easily noticable change will be visible if a small change in output in those input areas, resulting in the Vanishing Gradient.To solve this problem 2 solutions were used:

• Layer-by-layer Pre-Training

• Long Short-Term Memory (LSTM)

– A research paper was published in 2001 by Gartner in which they narrated the obstacles and opportunities of 3-D data growth. They explained that the range of data sources and types is also expanding as the data speed and data volume are growing. This was the call for preparation for Big Data's aggressive progress.

– ImageNet was introduced in 2009 by Fei-Fei Li, a professor at Stanford. An image database that is organized according to the hierarchical order of WordNet. With at least 14 million images, he assembled a database of labeled images. As we know, to train neural networks, we need labeled images, and the internet is filled with unlabeled images. "Data drives learning," Mr. Li said. Our vision was that Big Data would change the functioning of Machine learning.

– With the drastic increase in the speed of the GPU in 2011, Convolutionary Neural Network (CNN) training was possible without the use of layer-bylayer pre-training. And it was evident that Deep learning had an advantage in speed and efficiency as the speed of computing increased. One of the examples is AlexNet, a CNN designed in collaboration with Ilya Sutskever and Geoffrey Hinton by Alex Krizhevsky, whose architecture won international competitions such as ILSVRC (ImageNet Large Scale Visual Recognition Challenge) on 30 Sept 2012.

– In 2012, Google Brain published the conclusion of an abnormal project known as The Cat Experiment."This was a free-spirited project designed to explore the difficulties of Unsupervised Learning. While training Deep learning model we use supervised learning that means the data is labelled. To use unsupervised learning means we give unlabeled data to convolutional neural network and as to find recurring patterns. The cat experiment performed 70% better than the previous models, but it was able to recognize only 16% objects used in training and performed even more poorly while processing rotated or moved objects.

5. CURRENT SCENARIO OF DEEP LEARNING: Deep learning is all around us. Nowadays, Deep learning is used to determine which online advertisement to show in real time according to your browsing history, to identify your friend tagged in your Instagram or Facebook photo, translate your voice to text or to translate texts of a web page in different language. Deep learning is an important technology that's achieving results that seems to be impossible before. As a result of some recent advancements, Deep learning is performing better than humans in certain tasks like classifying images according to objects. Deep learning can also be found in intangible uses like Fraud detection by credit card companies, e-commerce companies use it to predict that how much people will unsubscribe their services and to provide personalized recommendations to customers, banks use Deep learning to predict bankruptcy and bad debts risks. Deep learning is being used widely to process automation, performance refinement, pattern detection and problem solving, and the range of Deep learning uses is limitless

Fig 5: Uses of Deep Learning.

6. RESULTS AND DISCUSSIONS:

6.1 Trends moving Deep learning to future

• Deep learning's current growths in research and industry applications show its omnipresence in every aspect of AI, be it Natural Language Processing or Computer Vision applications.

• If right amount of time and research opportunities will be provided, then unsupervised learning may deliver models that will closely imitate the behavior and thinking process of human.

• The obvious conflicts between consumer data protection laws and needs of large volumes of data for research purposes will continue.

• Limitations of Deep learning in being able to reason between good and bad, right and wrong is an obstacle while automating decision support tools.

• Though worldwide popularity, Deep learning may not be the sole savior of AI solutions.

• If Deep learning technology research progress will continue as per current speed, then soon developers may find themselves outpaced and will be forced to take in depth training.

• Machine learning and Deep learning must show learning from short amount of training materials, and transfer learning between context, continuous learning and adaptive capabilities to remain useful.

6.2 Feasibility of Deep learning

The future accomplishment of Deep learning depends on 2 factors. First is the availability of powerful computation power and the second is a large amount of data. In terms of computation power, we are growing much faster than what is described in Moore's law that is computational capability will grow 2 times every 2 year and with the wide acceptance of digital technologies, we are creating more data than ever before, making applied Deep learning more practical.

6.3 Future scenario of Deep learning

Current state of Deep learning is like infant's brain. And infants' brain is like a sponge, it will take some time for neural networks to be mature or reason like a grown-up human. But as the new neural network architectures is developing like Generative Adversarial Networks which can classification of images using one shot learning or OpenAI's GPT-2 Model which can create reasonable paragraphs of tax, reading paragraphs and summarizing text with human like accuracy, we are closing towards designing Deep-learning-based systems which can correctly imitate the complex functionalities of a human brain. In future Deep learning systems can exceed human intelligence, that will result in advance level cognitive system that can end fluently interact with humans.

7. CONCLUSION

All though Deep learning is growing faster than ever, but there are still ways to go. We are still far away from fully understanding that how human brain works, how we can make machine smarter, close 2 are smarter than humans or how can we design system that can imitate the learning process of human brain. No doubts, Deep learning has been providing solution for many problems while helping in developments of training to another level. Even though the multitude of open research issues and the fact at the Deep learning is still in its infancy, the development made in developing deep Machine learning system will shape the future of Machine learning.

REFERENCES

[1] Matiur Rahman Minar, Jibon Naher "Recent Advances in Deep learning: An Overview", 2016, Ben-Bright Benuwa, Yongzhao Zhan, Benjamin Ghansah, Dickson Keddy Wornyo and Frank Banaseka Kataka.

[2] D. Yu and D. Li, 'Deep learning - Methods and Applications,' Foundations and Trends in Signal Processing, Now Publishers, 2014.Gyusoo Kim and Seulgi Lee, "2014 Payment Research", Bank of Korea, Vol. 2015, No. 1, Jan. 2015.

[3] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, pp. 436-444,

2015.Hitesh D. Bambhava, Prof. Jayeshkumar Pitroda, Prof. Jaydev J. Bhavsar (2013), "A Comparative Study on Bamboo Scaffolding And Metal Scaffolding in Construction Industry Using Statistical Methods", International Journal of Engineering Trends and Technology (IJETT) – Volume 4, Issue 6, June 2013, Pg.2330-2337.

[4] Y. LeCunn "Deep learning: Past, Present and Future", Facebook AI Research, New York University, 2018.

[5] Hal S. Greenwald, Carsten K. Oertel, "Future Directions in Machine learning". Frontiers in robotics and AI, 2017.

[6] Gilad David Maayan, "Deep learning: Past and Future", Datafloq.com, 2019.

[7] Anandhavalli Muniasamy, Areej Alasiry, "Deep learning: The Impact on Future eLearning", Vol. 15, 2020