Review Article **Open Access**

Deep Learning – An Overview

Adedokun OG*

Department of Electronic Science and Technology, University of Science and Technology China, Hefei, China

Abstract

In recent years deep learning has become the big thing happening in the field of Machine learning with so many research and discovery in the field, within this few period deep has had a big edge over other forms of machine language since it has made a better attempt at learning a big amount of unlabeled data and it has been applied to so many fields. The recent success in deep learning has provided significant contribution in the field of artificial intelligence. This article presents a review on deep learning, basic overview of history of deep learning, and a quick overview of some important concept of deep learning (how feature learning made the difference in deep learning and the key concept of deep learning which the big data is). In a quick summary this article presents a basic overview of deep learning and a track of its progress.

Keywords: Deep learning; Machine learning; Neural network

Introduction

Machine Learning is a subfield of Artificial Intelligence which involves providing machines with the data they need to learn for the machine to do something or make decisions without being explicitly programmed to do it. Algorithms such as decision tree learning, inductive logic programming, clustering, reinforcement learning, or Bayesian networks helps them make sense of the inputted data. Machine learning was a giant step forward for Artificial Intelligence.

The development of neural networks – a computer system set up to classify and organize data much like the human brain – has advanced things even further. Based on this categorization and analysis, a machine learning system can make an educated "guess" based on the greatest probability, and many are even able to learn from their mistakes, making them "smarter" as they go along. But in recent year the birth of deep learning has advanced things further and they have produced results comparable to and in some cases superior to human experts. Just as Machine learning became a giant step forward for Artificial Intelligence Deep learning is the new giant step in machine learning.

Deep Learning is a subfield of machine learning which employs algorithms to process data and imitates the thinking process to develop premonition. *It uses the machine learning methods based on learning data representation.* Deep Learning uses layers of algorithms to process data, understand human speech, and visually recognize objects. It's a sub of machine learning artificial which is neural network inspired.

According to [1] deep learning is making a good wave in delivering solution to difficult problems that has been faced in the field of Artificial Intelligence for so many years, and also has had a landmark success in an all-round field of science, with emphasis on Image and sound processing including facial recognition, speech recognition, computer vision, and many others. According to their publication in 2015, they stated the fact that it has beaten other machine-learning techniques in all round aspect of Artificial intelligence.

It can be simply put as a set of learning method that model data with complex architecture, having the structure and the function of the brain, the neural network as its elementary bricks. Neural network of multiple layers are combined to form the deep neural network (the more the network the deeper the network is said to be).

There exist several types of architectures for neural networks:

- The multilayer perceptron, they are the oldest and simplest ones.

- The Convolutional Neural Networks (CNN), CNN is particularly adapted for image processing.

- The recurrent neural networks used for sequential data such as text or times series.

History of Deep Learning

Neural networks are composed of layers of many simple processors called neurons, with connections in different layers. These networks transform data until they can classify it as an output. Neurons get activated through weighted connections from previously active neurons Figure 1 [2].
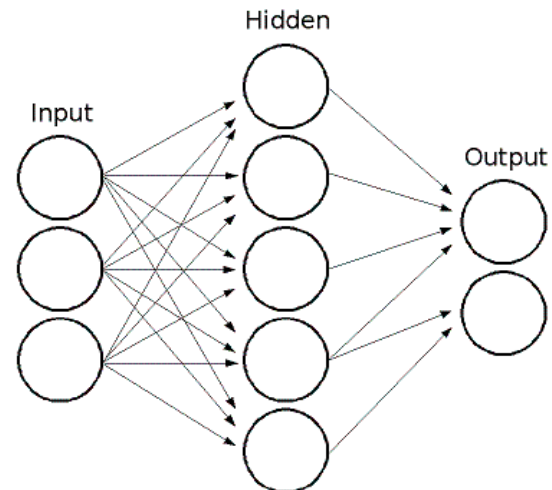

Figure 1: Neural Networks.

*Corresponding author: Adedokun OG, Department of Electronic Science and Technology, University of Science and Technology China, 1129 Huizhou Ave, Baohe Qu, Hefei Shi, Anhui Sheng, China, 230022, Tel: +8615555455502; E-mail: toba@mail.ustc.edu.cn

Iterative learning process is a key feature of neural network in which records are presented to the network in tandem in respect to associated weight that makes the neural network exhibit required behavior, it may require long chain of computational stages depending on the prediction to be achieved.

Neural network has high tolerance to noisy data and can classify pattern on which they've not been trained [2,3].

Neural network models with few stages have been around for many decades, Models with several successive nonlinear layers of neurons dated back to as long as 1960s An efficient gradient descent method for teacher-based Supervised Learning in discrete, differentiable networks of arbitrary depth called back propagation.

Back propagation is the most popular neural network algorithm it was developed in the 1960s and 1970s, In the year 1986 deep learning was introduced by Rina Dechter to the machine learning community [2] and in 2000 to artificial neural network in by Igor Aizen berg [4-6] in 1989 applied back propagation to neural network in a paper termed Back propagation Applied to Handwritten Zip Code Recognition. Back propagation based training of deep neural network with many layers, was researched more on in the 90's because it didn't experience so much progress in the 80's Deep learning became practically feasible to some extent through the help of Unsupervised Learning in 1991 [7] and made visible progress in 2006, Geoffrey Hinton a pioneer in the field of artificial neural networks in 2006 co-authored a paper in which they describe an approach to training many layered network of restricted Boltzmann machines [8] it was well received in the academic community a successful example of greedy layer network, he never stopped there, he still further in the same area same year while sticking with the same layer network idea "deep" greedy layered network [9].

The 1990s and 2000s also saw many improvements of purely supervised Deep Learning featuring works from German computer scientist Schmidhuber, Support Vector Machine SVMs a system for mapping similar data, recognize characters, classify images as it relates to machine learning tweaked and refined by Cortes and Vapnik in 1994 [10], 1998 Yann LeCun made another contribution in the field of deep learning with his publication on gradient based learning [11]. The year 2009 gave birth to Image Net, a large and free database of labelled images available to researcher which was launched by Li Fei fei [12]. Alex Krizhevsky in 2012 created the AlexNet a convolutional neural network its success kicked off a convolutional neural network renaissance in the deep learning community [13].

2014 experienced a big leap in deep learning, the creation of Deep face. Deep face was developed and released to the world in 2014, it's a deep learning facial recognition system created by facebook nicknamed deep face it use neural network to identify faces with about 97.35% accuracy over 13% more than FBI's Next generation Identification system [14].

Deep Learning is Large Neural Network

Andrew Yan-Tak Ng, Chief scientist at Baidu who Lead Google Brain a deep learning artificial intelligence research formed in 2010, In the 2013 talk titled "Deep Learning, Self-Taught Learning and Unsupervised Feature Learning" and 2015 talk at ExtractConf2015 "What data scientists should know about deep learning" he pointed it out that the core of deep learning is enough data, according to him he perceive we can leverage on the availability of fast enough computers capable of handling lots of data to train large neural network, in his talk in 2015he emphasized the fact that it's all about scaling. We construct large neural network and train them with more data, this will lead to increase in their performance. The picture below is an extract from his talk in the Extract Conf 2015 which explains the big data approach to deep learning Figure 2.

In 2016 Jeff Dean in his talk titled "Deep Learning for Building Intelligent Computer Systems" in the same attitude emphasized that deep learning is all about large neural networks. He mentioned that deep learning is deep Neural network. Figure 3 below is a picture of Jeff Deans slide [15].

Feature Learning

The ability of deep learning model to perform automatic feature extraction from raw data is called feature learning. Yoshua Bengio described deep learning as the ability of the algorithms to discover and learn good representation using feature learning [16].

Feature engineering is the process of creating features that makes machine algorithm works by using domain knowledge of data. Feature engineering is key to the application of machine learning [17], it relies on human domain knowledge much more than data, and also if handcrafted features have multiple parameters it makes it hard to manually tune them in all developing effective features for new application is slow, difficult and expensive. The need for manual feature engineering can be obviated by automated feature learning.

Feature Learning is the ability of deep learning model to perform automatic feature extraction from raw data is called feature learning. Yoshua Bengio described deep learning as the ability of the algorithms to discover and learn good representation using feature learning [16]. The feature learning technique lets a system to easily extract useful
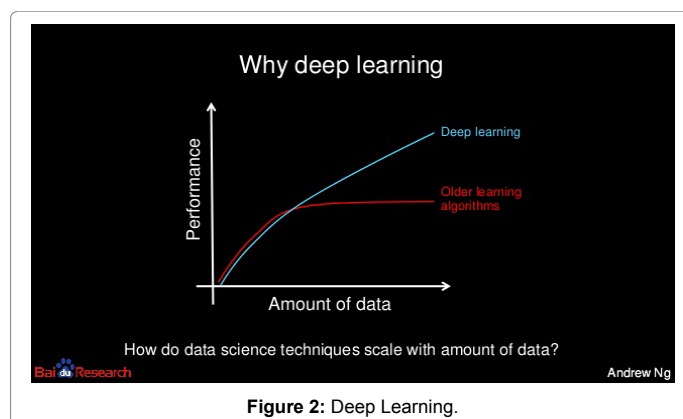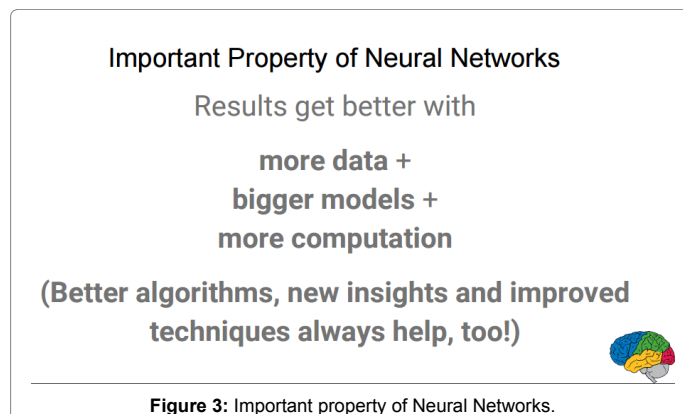

Figure 2: Deep Learning.

Figure 3: Important property of Neural Networks.

information needed for feature detection or classification from raw data. The feature learning replaces manual feature engineering. The feature learns the value of a huge number of parameters in feature representations; it's faster to get feature representations for new application therefore makes better use of big data. Deep learning means feature learning, in Deep learning Feature learning makes the difference. Featured Learning is categorized into supervised learning or unsupervised learning.

Unsupervised learning

Unsupervised learning is a type of machine learning algorithm that learns from datasets consisting of input data that has not been labeled, it would identify patterns, anomalies and similarities in the data. It makes the data more readable and organized. In unsupervised learning, the model is given a dataset which is neither labelled nor classified there are no label for training data. The model explores the data and makes deductions from data to define hidden structures from unlabeled data. Although it cannot add label clusters, for example it cannot say this group of cats or dogs, but it will separate all the cats from dogs.

Cluster analysis is the most common unsupervised learning method, it is used to find hidden patterns or grouping in data it runs through the presented data and finds these natural clusters. It groups a set of objects of same type into a group called a cluster and more similar object types into another groups (clusters) a good example is the Google's cat detector model [18]. Some of the clustering algorithm includes but not limited to Hierarchical clustering, *Probabilistic Clustering*, and k-Means clustering.

Application

Deep learning can be broadly utilized in numerous areas of life and applications and fields which includes but not limited to automatic speech recognition, Image recognition, visual art processing, natural language processing, drug discovery and toxicology, customer relationship management, bioinformatics, medicine and so many more. For example in the medicine field, deep learning had been used to predict quality of sleep based on data from wearable's [19], also for image recognition, recent years deep learning has proven to produce more accurate result than human in this filed [20] for image recognition it employs CNN Convolutional Neural Networks which uses relatively little pre-processing compared to other image classification algorithms. It expects and preserves the spatial relationship between pixels by learning internal feature representations using small squares of input data. Feature are learned and used across the whole image, allowing for the objects in the images to be shifted or translated in the scene and still detectable by the network. It is this reason why the network is so useful for object recognition in photographs, picking out digits, faces, objects and so on with varying orientation. One of the advantages of CNN is that they automatically learn and generalize features from input domain.

Conclusion

This article has presented deep learning and has shown that the core of deep learning is big data as presented by Andrew Ng who emphasized how advantage of fast computers can be taken to run large data. Deep learning is a big neural network with a lot more data,

requiring bigger computers. Although early approaches published by Hinton and collaborators focus on greedy layer wise training. Deep learning has achieved a quick growth in the field of machine learning the basic historical overview of deep learning done in this study has shown the importance of this field and has demonstrated the growth of deep learning.

This article has been able to show a review on deep learning, basic overview of history of deep learning, and a quick overview of some important concept of deep learning. The field of deep learning is a wide one thus, the later part of this study has been narrowed down to some selected specifics of deep learning and has only discussed one of the many application.

Reference

1. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521: 436-444.

2. Schmidhuber J (2015) Deep learning in Neural Networks: An overview. Int J Appl Eng Res 10: 25433-25448.

3. Shah J. An introduction to neural networks learning.

4. Gomez FJ, Schmidhuber J (2005) Co-evolving recurrent neurons learn deep memory POMDPs. GECCO Proc, Washington DC, USA, PP: 491-498.

5. Beyer HG, O'Reilly UM (2005) Genetic and Evolutionary Computation Conference, (GECCO) 2005, Proceedings, Washington DC, USA.

6. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, et al. (1989) Back propagation Applied to Handwritten Zip Code Recognition. Neural Comput 1: 541-551.

7. Schmidhuber J (1992) Learning Complex, Extended Sequences Using the Principle of History Compression. Neural Comput 4: 234-242.

8. Hinton GE, Osindero S, Teh YW (2006) A Fast Learning Algorithm for Deep Belief Nets. Neural Comput 18: 1527-1554.

9. Hinton G, Salakhutdinov RR (2006) Reducing the Dimensionality of Data with Neural Networks. Science 313: 504-507.

10. Drucker H, Cortes C, Jackel LD, LeCun Y, Vapnik V (1994) Boosting and Other Ensemble Methods. Neural Comput 6: 1289-1301.

11. Lecun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc. IEEE 86: 2278-2324.

12. Deng J, Dong W, Socher R, Li L, Li K, et al. (2009) ImageNet: A large-scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition Pp: 248-255.

13. Krizhevsky A, Sutskever I, Hinton GE (2012) Image Net Classification with Deep Convolutional Neural Networks. ACM DL 60: 84-90.

14. Taigman Y, Yang M, Ranzato M, Wolf L (2014) DeepFace: Closing the Gap to Human-Level Performance in Face Verification. CVPR Pp: 1701-1708.

15. Dean J (2016) Large-Scale Deep Learning for Building Intelligent Computer Systems. BT - Proceedings of the Ninth ACM International Conference on Web Search and Data Mining, San Francisco, CA, USA.

16. Bengio Y (2012) Deep Learning of Representations for Unsupervised and Transfer Learning. Proceedings of ICML Workshop on Unsupervised and Transfer Learning Pp: 17-36.

17. Ng A (2013) Machine Learning and AI via Brain simulations. This Talk.

18. Le QV, Ranzato MA, Monga R, Matthieu Devin, Kai Che (2011) Building high-level features using large scale unsupervised learning. Proc ICML 1: 507-514.

19. Sathyanarayana A, Joty S, Fernandez-Luque L, Ofli F, Srivastava J, et al. (2016) Sleep Quality Prediction From Wearable Data Using Deep Learning. JMIR mHealth uHealth 4: e125-e125.

20. Cireşan D, Meier U, Masci J, Schmidhuber J (2012) Multi-column deep neural network for traffic sign classification. Neural Networks 32: 333-338.