International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 07 | July 2024 www.irjet.net p-ISSN: 2395-0072

An Overview of Deep Learning: Milestones, Models and Trends

Sheela N S[1], Surekha Bijapur[1], Dr. J Venkata Krishna[2]

[1] Research Scholar, Dept. of Computer Science and Engineering, Srinivas University, Mangaluru, Karnataka, India
[2] Associate Professor, Dept. of Computer Science and Engineering, Srinivas University, Mangaluru, Karnataka, India

---

Abstract - *Neural Networks are inspired by human brain, to make the computer to process data. Neural Networks are used in Deep Learning which is a subset of Machine Learning and Artificial Intelligence. Many models of Neural Networks are implemented based on the target applications. This paper provides an overview about deep learning. Initially, the development of deep learning is described. Later it focuses on the basic concepts of Neural Networks essential to understand the working. Thirdly it focuses on few key models such as Feed Forward Neural Network, Recurrent Neural Network, and Convolutional Neural network their architecture and working. At last, it discusses about the trends of deep learning.*

Keywords: **Deep Learning, Neural Networks, Artificial Intelligence, Machine Learning, Feed Forward Neural Network, Convolutional Neural Network, Recurrent Neural Network.**

1. INTRODUCTION

Deep learning, a class of machine learning techniques teaches the computers to process data in a way the human brain does. Deep learning uses multilayered neural networks which is inspired by human brain to solve the complex problems. The key difference between deep learning and machine learning is the underlying neural network architecture. Machine learning uses simple neural network whereas Deep learning uses more layered neural networks.

2. HISTORICAL DEVELOPMENT

The evolution of deep learning can be traced back to the early days of artificial intelligence and neural networks. The primary goal of deep learning and artificial neural networks is to make a computer system to simulate human brain. The history of deep learning takes us back to 300 BC, Associationism theory stated by Aristotle [1]. Associationism is a theory states that mind is a set of conceptual elements that are organized as associations between these elements [2].

The evolution of deep learning starts in 1942, with the concept of artificial neuron. Warren McCulloch and Walter Pitts developed a mathematical model which based on the working of basic biological neuron. That artificial neuron is called MCP-neuron. That laid the foundation for Deep Learning [2].

In 1949 Donald Hebb considered as the father of neural networks, introduced Hebbian Learning Rule, in his book "**The Organization of Behavior**" which lays the foundation of modern neural network. This rule explains how neurons adapt and form stronger connections through repeated use [3]. In 1958 Frank Rosenblatt introduced the first perceptron to mimic the human brain's learning process, which highly resembles modern perceptron [4].

Back-propagation a learning procedure which repeatedly adjusts the weights of the connections in the network to minimize a measure of the difference between the actual output vector of the net and the desired output vector is proposed by David E. Rumelhart et al. in 1986 [5].

In 1998 LeNet-5 Developed by LeCun et al., LeNet was one of the first successful applications of convolutional neural networks (CNNs), a Gradient-Based Learning Applied to handwritten digit recognition tasks was a major step in learning from data [6].

Hinton et al. proposed deep belief networks (DBNs) in 2006, a generative model for digit classification. This reignited interest in deep learning by demonstrating their capability to learn hierarchical representations [7]. In 2012 Krizhevsky et al.'s AlexNet achieved significant breakthroughs using deep convolutional neural network to classify high-resolution images classification [8], marking the beginning of the deep learning era.

Rob Fergus et al. made an improvement of AlextNet with greater accuracy and named as ZFNet [13] in 2013. A GoogleNet [19] was propsed by Szegedy et al. in 2014 called as Inception-V1 to assess the quality in the context of object detection and classification. In 2014 VGGNet [20] a deep convolutional network introduced by Simonyan et al. for large-scale image classification. Later in 2015 Kaiming He et al. used Residual Networks (ResNets) [21] for computer vision applications like object detection and image segmentation.

3. CONCEPTS OF NEURAL NETWORKS

In this section we are discussing very essential concepts and terminologies in order to understand deep learning [9].

Neuron: a neuron forms the basic structure of a neural network. A neuron is a mathematical model receives an

input, computes the weighted average of its input and then applies a bias to it. Post that, it passes this resultant term through an activation function and generates a nonlinear output which is either sent to other neurons for further processing or it is the final output.

Activation function: it is mathematical function which takes linear input and produce nonlinear output. There are three types of activation function:

Binary Step: It is a very simple activation function. The binary output is produced by the function which is based on whether the input is above or below certain threshold.

Linear activation function: It is also a simple activation function. It is similar to straight line equation f(x)=x. The output is proportional to input. It is used only in the output layer as they don't capture nonlinear data.

Nonlinear activation function: It is a very complex activation function. It produces nonlinear output. These functions are used to add non linearity, through which it is possible to achieve non-linear mappings from inputs to outputs. Thus it is important to use non-linear activation functions in neural networks [10].

The most commonly used non-linear activations are:

Sigmoid or Logistic Activation Function: It is the common activation function used in neural network [11]. It is given as below:

$$f(x) = \frac{1}{1+e^{-x}}$$

in which, $x \in (-\infty, +\infty)$, $f \in (0,1)$

Tanh Function (Hyperbolic Tangent): It is same as sigmoid activation function. It ranges between -1 to 1. The equation of tanh function is:

$$f(x) = \frac{(e^{x} - e^{-x})}{(e^{x} + e^{-x})}$$

ReLU Function: Rectified Linear Unit is also one of the most commonly used activation function in neural network. It is a most efficient function. It is mathematically represented as:

$$f(x) = max\,(0, x)$$

The above functions are the most commonly used activation functions. There are the variations of the functions which are used as activation functions.

Weights: Weights are the numerical values associated with each input values. It determines the importance of the input in predicting the input. Initially the weights are selected randomly and later adjusted in the back propagation.

Bias: It is a numeric value which help us to shift the activation function. There are many techniques to initialize the bias value.

Epoch: An epoch is one cycle of training the neural network with all the training data. This one cycle includes both forward pass and backward pass.

4. WORKING OF NEURAL NETWORK

Neural network makes the decisions in a manner similar to the human brain. Figure 1 [22] shows the working of Neural NetworkThe input is passed through the network layers to generate output. The layers involved in the model are one input layer, one output layer and one or more hidden layers. The input is fed into input layer. The hidden layers process the input data by applying the activation function to the weighted sum of these inputs and passes the result to next layer. This will continue until it reaches the output layer. In output layer the network output is generated. The loss score is calculated and optimizer is used to adjust the weights and biases in backpropagation.

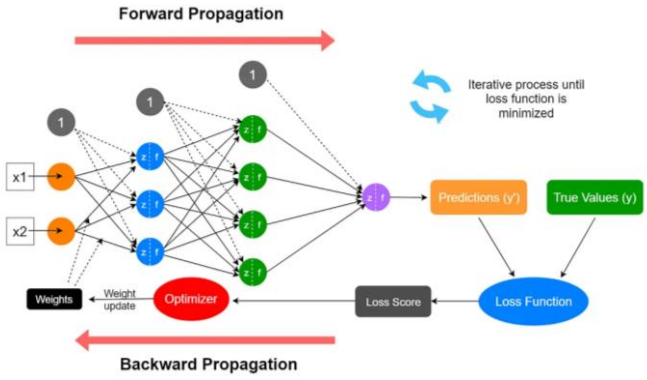

Figure 1: Working of Neural Network

5. KEY MODELS

In this section we are focusing on more popular neural network models such as Feed Forward Neural Networks, Recurrent Neural Networks and Convolutional Neural Networks amongst the available models. We are presenting the architecture, working and applications of those models.

5.1. Feed Forward Neural Networks

Feed forward neural network is very simple and basic neural networks. As the name indicates the input flows from input layer to output layer through hidden layers. There is no feedback from the output layer to input layer. The model compares predicted output with the actual output and adjust the weights according to the error value. Many such iterations result in the more accurate prediction of output.

Feed forward neural networks are categorized into two categories single layer and multi-layer based on the layers [12].

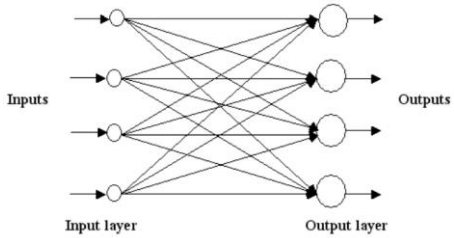

Figure 2. A single layer feed forward neural network

Single layer feed forward neural network is as show in the figure 2 [12]. It has only two layers input and output. Each input is given to output layer directly to produce output. The final output is computed by summing up the outputs from all neurons of output layer. This model is limited in handling complex data because of single layer.

Multi-layered feed forward neural network is also called as multi-layer perceptron (MLP). It has a multiple layers of neurons. The layers are an input layer, an output layer and one or more hidden layers. Figure 3 [12] shows one of such MLP. It is referred as 5-3-2 network because there are 5 input neurons, 3 hidden neurons and 2 output neurons.

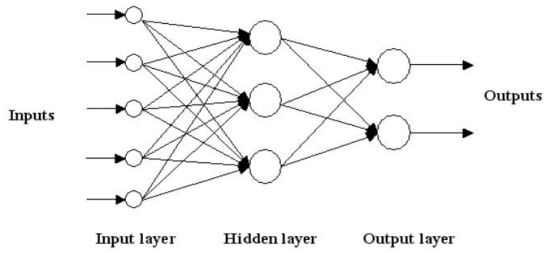

Figure 3. A multi-layer feed forward neural network

Back Propagation Algorithm

Back propagation algorithm is one of the most popular algorithms used to train the feed forward neural networks. It is a supervised learning method the weights of the neural network are adjusted using this algorithm. During a learning phase, the output of the network is compared with the actual output to know how well the model is working based on the difference. This difference is called as error value based on this value the weights are adjusted in each epoch of the learning phase.

5.2. Convolutional Neural Networks (CNNs)

CNNs are specialized neural networks also known as ConvNet. It is basically used for computer vision tasks like object recognition, image classification, detection and segmentation.

Components

The layers such as convolutional layers, pooling layers and fully connected layers the keys components of convolutional neural networks. A typical CNN architecture is as shown in the figure 4 [23]. It consists of repetitions of a stack of several convolution layers and a pooling layer, followed by one or more fully connected layers.

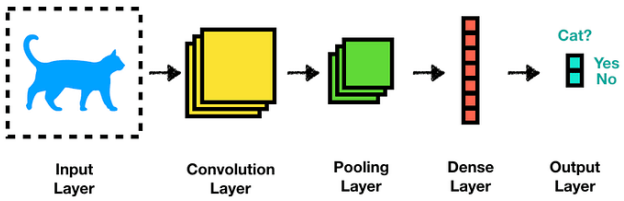

Figure 4: A CNN applied to object recognition

Convolution layer

It is the fundamental component in CNN the mathematical operation convolution is performed. Convolution is a linear operation used to extract the feature. The steps of this operation are as follow.

1 Kernel a small array of numbers is applied across the tensor an array of numbers that represents the input.

2 Element-wise multiplications are performed between each element of kernel and the tensor at each location of the tensor starting from the top-left corner to right.

3 The product from the above step is summed up to obtain output value called as feature map.

4 The above steps are repeated until the image matrix is fully covered.

The convolution operation is as shown in the Figure 5 [14]. The output of convolution operation is then passed to the nonlinear activation function ReLU.

Pooling layer

Pooling layer provide down sampling operation, which is a technique used to reduce the dimensions of feature maps while retaining the essential information. There are many types of pooling [26]

- Max pooling: will select the maximum value of the feature map region covered by the filter.
- Global pooling: Every channel of the feature map gets reduced to single value.
- Average pooling: The average value of the presented elements in the feature map's region covered by the filter is computed.

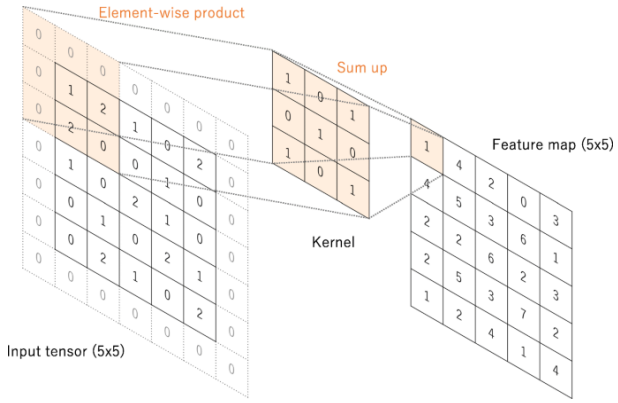

Figure 5: Convolution Operation

Fully connected layer

It is also known as dense layer [14]. The input to this layer is one-dimensional matrix generated by the last pooling layer. The ReLU is the activation function used in this layer.

5.3. Recurrent Neural Networks (RNNs)

RNNs are another key model of Neural Networks. They are derived from feed forward neural networks but internal memory capability makes it different from the feed forward neural network (figure 6 [25]). RNNs are used for sequence prediction. They are also popular in natural language processing.

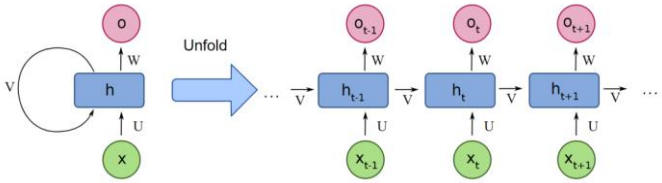

Figure 6: Recurrent Neural Network

Figure 7 [24] shows the types of RNN. There are four types of RNN based on the number of inputs and number of outputs in the network such as:

- One-to-one: one input and one output
- One-to-many: one input and many outputs
- many-to-one: many inputs and one output
- many-to-many: many inputs and many outputs

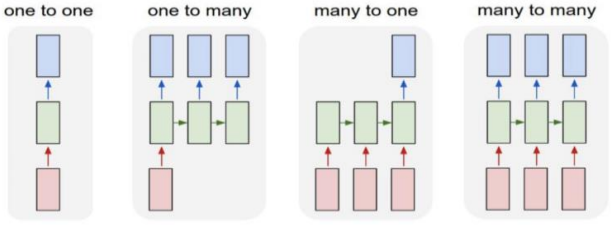

Figure 7: Types of RNN

6. TRENDS OF DEEP LEARNING

Generative AI- Generative AI is an AI technology that can generate various new contents like text, image, audio etc., based on what it has learned from existing content [15]. It is a subset of Deep Learning make use of Generative Adversarial Networks and implement the models.

Natural Language Processing(NLP)- NLP is a sub field of AI uses the machine learning and deep learning to make a device to understand and generate the text [16]. NLP is used wide range of applications in daily life like search engines, chatbots, virtual assistance and much more.

Explainable AI(XAI)- XAI is an AI which produce output and make users to understand it easily by providing transparency of the model working. This makes the model trust worthy. The application domains of XAI are agriculture, healthcare, finance, computer vision and so many [17].

Continual Learning- Continual Learning (also known as Incremental Learning**,** Life-long Learning) **is a** concept to learn a model for a large number of tasks sequentially without forgetting knowledge obtained from the preceding tasks, where the data in the old tasks are not available anymore during training new ones. It has three scenarios task-incremental learning(Task-IL), domain-incremental Learning (Domain-IL) and class-incremental learning(Class-IL) [18].

Hybrid models- This model combines different types of deep neural networks or other ML techniques in order to get the good performance and accuracy.

7. CONCLUSION

Deep learning is a subset of Artificial Intelligence provides powerful tools for data analysis and decision-making as result revolutionized many fields. In This review paper we highlighted the historical development, key architectures and trends of deep learning. Along with continuous evolution and addressing current challenges it provides us the future research directions.

REFERENCES

[1] WH Burnham. Memory, historically and experimentally considered. i. an historical sketch of the older conceptions of memory. The American Journal of Psychology, 2(1):39–90, 1888.

[2] Haohan Wang & Bhiksha Raj “On the Origin of Deep Learning”, arXiv:1702.07800 **[cs.LG],** 2017

[3] D.O. Hebb, The organization of behavior; a neuropsycholocigal theory, A Wiley Book in Clinical Psychology 62 (1949) 78.

[4] Rosenblatt, F. (1958). The Perceptron: A probabilistic model for information storage and organization in the brain. Psychological Review, 65(6), 386-408.

[5] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533-536.

[6] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.

[7] Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. Neural Computation, 18(7), 1527-1554.

[8] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 1097-1105.

[9] Bishop, Christopher M., and Hugh Bishop. *Deep learning: Foundations and concepts*. Springer Nature, 2023.

[10] Sharma, S., Sharma, S., & Athaiya, A. "Activation functions in neural networks", International Journal of Engineering Applied Sciences and Technology, 2020 Vol. 4, Issue 12, ISSN No. 2455-2143, Pages 310-316

[11] B. Ding, H. Qian and J. Zhou, "Activation functions and their characteristics in deep neural networks," *2018 Chinese Control And Decision Conference (CCDC)*, Shenyang, China, 2018, pp. 1836-1841,

[12] Sazlı, M. H. (2006). A brief review of feed-forward neural networks. *Communications Faculty of Sciences University of Ankara Series A2-A3 Physical Sciences and Engineering*, *50*(01).

[13] Zeiler, M. D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I 13* (pp. 818-833). Springer International Publishing.

[14] Yamashita, R., Nishio, M., Do, R.K.G. *et al.* Convolutional neural networks: an overview and application in radiology. *Insights Imaging* **9**, 611–629 (2018).

[15] Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative ai. *Business & Information Systems Engineering*, *66*(1), 111-126.

[16] Khurana, D., Koli, A., Khatter, K. *et al.* Natural language processing: state of the art, current trends and challenges. *Multimed Tools Appl* **82**, 3713–3744 (2023).

[17] Saranya A., Subhashini R., A systematic review of Explainable Artificial Intelligence models and applications: Recent developments and future trends, Decision Analytics Journal, Volume 7, 2023, 100230, ISSN 2772-6622

[18] Van de Ven, G. M., & Tolias, A. S. (2019). Three scenarios for continual learning. *arXiv preprint arXiv:1904.07734*.

[19] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9).

[20] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

[21] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Identity mappings in deep residual networks. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14* (pp. 630-645). Springer International Publishing.

[22] https://medium.com/data-science-365/overview-of-a-neural-networks-learning-process-61690a502fa

[23] https://towardsdatascience.com/convolutional-neural-network-a-step-by-step-guide-a8b4c88d6943

[24] https://builtin.com/data-science/recurrent-neural-networks-and-lstm

[25] https://www.analyticsvidhya.com/blog/2022/03/a-brief-overview-of-recurrent-neural-networks-rnn/

[26] https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/