

ARTIFICIAL INTELLIGENCE AND DEEP LEARNING METHODOLOGIES


B. Alafi

Abstract— In this paper, a brief definition and history of Artificial Intelligence and deep learning are mentioned. The procedure from Artificial intelligence to deep learning is described. Due to human loss and disabilities in complicated problems and huge data amount, Artificial Intelligence and learning methods were introduced. Their main aim is to make machines perform like humans. These two – Artificial Intelligence and Deep Learning – are not separated from each other, but their application is different. As a sequence, the main concept of the algorithm for each one is announced and compared. In the end, all remarks on Artificial Intelligence and Deep learning are discussed.

Keywords— *Artificial Intelligence, Deep learning, AI, Learning system.*

1. INTRODUCTION

THE words Artificial Intelligence was first used in 1956, and has become very popular due to advanced algorithms such as increased data volumes, advances in computing power and storage AI. due to human disability in facing with some computing and storage issues. This topic was welcomed by defense agencies for some topics like problem-solving, symbolic methods, mapping projects. As a result, the first individual assistant was introduced during the first street mapping project in 1970. While in Hollywood films including, Transformers AI is depicted as human-like robots that take over the world [1,2]. The latest evolution of AI technology, however, is neither that frightening nor that wise. Alternatively, AI has benefited from various sectors, including modern examples such as healthcare, banking, fashion, education, and more. AI's main goal is to give machines human intelligence. This focuses specifically on making the machines smarter and thinking as well as behaving like humans. These machines are being trained for problem-solving and thinking better than humans do. The best examples of AI are self-driving cars and robots [5,6]. Regarding the mentioned topics, Machine Learning is an AI sub-set which focuses exclusively on making predictions based on experiences. It allows the machine to make a decision based on knowledge and data rather than an explicit program to perform a specific task. The algorithms are specifically designed to learn and improve over time and help the user make a better choice [5]. Machine Learning is divided to three types: supervised, unsupervised, and reinforcement or semi-supervised learning [7]. Different issues especially on data subject – analyze data, Collect data, Filter data, Train algorithms, Test algorithms, Use algorithms for future predictions – are dealt with Machine Learning; image

Behnaz Alafi, is with Electrical Engineering Faculty Istanbul Technical University, Istanbul, Turkey, (e-mail: alafi18@itu.edu.tr). 

Manuscript received Sept 09, 2019; accepted Jul 11, 2019.
Digital Object Identifier:

recognition can be mentioned as an example for applying these methodologies. At last, Deep Learning is one of Machine Learning techniques that is recognized via neural networks and inspired by brain neurons performance. Deep neural networks which are inspired by the human brain's function are used by a collection of the task-specific algorithm as a neural network. The main inspiration of deep learning is by neuroscience which is affected by ongoing knowledge, observing, experience, tentative discovering. The classification of deep learning can be identified by patterns that provide expected output and result for a received input. Deep learning as an approach of machine learning is associated by machine learning but not interchangeable with limitations of it. At a far glance of these subjects, it is observed that deep learning is located inside the machine learning, and machine learning is located inside the Artificial intelligence, on the other word artificial intelligence is overarching all learning algorithms.

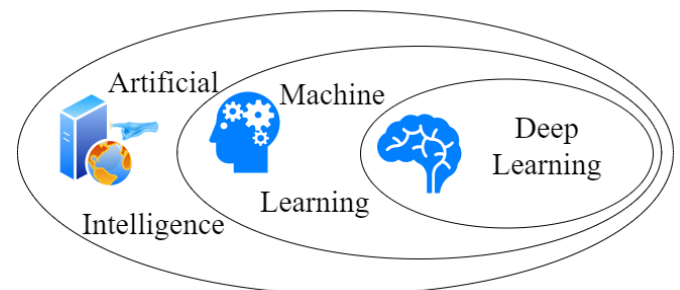


Fig.1. AI, Machine Learning, and Deep Learning interconnection

2. GENERAL CONCEPT OF ARTIFICIAL INTELLIGENCE (AI)

Artificial Intelligence is a branch of computer science that tries to copy or simulate human intelligence on a machine, so machines can often perform tasks that require human brain and abilities and totally intelligence. planning, learning, reasoning, problem-solving and decision making are considered as some programmable functions of AI systems. Algorithms via techniques like machine learning, deep learning, and relevant rules support Artificial intelligence systems. Learning and especially machine learning algorithms feed data to Artificial Intelligence systems and train them through statistical techniques. Machine learning helps AI systems to act better without the need for special planning.

2.1. KNOWLEDGE-BASED SYSTEM

A knowledge-based system is a kind of Artificial Intelligence system with the aim of capturing knowledge and experiences from human for problem-solving and decision making. Human learning, thinking, decision-making, and action are supported by knowledge-based techniques for a knowledge-based system.

Knowledge-based system is one type of Artificial Intelligent systems that its aim is taking human knowledge and expert to be trained. The quality of this kind of cooperating with human is related to aforesaid training process. So, utilized learning techniques and algorithms are important factor in knowledge-based systems.

2.2. MACHINE LEARNING

Machine learning is a technique which can support an Artificial Intelligent system to be trained without being programmed directly and clearly. Before the invention of learning algorithms, it was necessary to program the system clearly. It was very hard for complicated duties and systems and did not apply to all type of functions. By the way, learning techniques have been used for supporting systems by capturing knowledge and experiences and use them for training themselves. Machine learning uses computer programming to take data as its feed and use them for learning itself without human intervention. A machine learning algorithm is divided into two main categories: supervised, and unsupervised [6].

2.3. FUZZY LOGIC

Creation of Intelligent machines, and the effect of computer logic on creating these machines, influence their logic and make difficult to imitate the human behavior. To achieve the purpose of learning human behavior and imitate the way of human thinking the fuzzy theories and logic were created. In 1965, the first research paper about fuzzy sets was published by Lotfi A.Zadeh. fuzzy logic deals with real information which is uncertain, unclear, implicit, and without plain boundaries. On the other words, fuzzy logic is located in contrast to Boolean logic, it is considered as many-valued logic and involving any real number between 0 and 1 both inclusive. This kind of theories and logic provides a reasonable means for better analyze of objects. All fuzzy logic-based computational algorithms have been used in the construction of the intelligent system, and development of decision-making, control process, optimization, and pattern and image recognition [2].

3. DEEP LEARNING ALGORITHM(DLA)

Deep learning is a sub-part of the machine learning technique [5]. Since it can learn from an enormous amount of unstructured and unlabeled data – known as big data – under the unsupervised category, is known as a deep neural network or deep neural learning. Today, because of the data explosion from each area in the world that is called big data, deep learning has completed gradually in response to this data explosion. Gathering and processing of big data that is related to alternative sources in the world would take the human being a lot of times maybe decades. This vast amount of unstructured and unlabeled data is learned by deep learning. A brilliant example of deep learning implementation is on detection of fraud or money laundry. Deep learning consists of multiple simple technologies. Improvement of these technologies eventually developed deep learning. The most known algorithms of deep learning are the convolutional neural network (CNN), recurrent neural network (RNN), denoising autoencoder (DAE), deep belief networks (DBNs), long short-term memory (LSTM) [6-9]. CNN is one of the most popular

models of deep learning which is introduced by details in following sections as an instance.

3.1. MATHEMATICAL DESCRIPTION

As aforesaid, deep learning is a subset of machine learning which executes machine learning process via utilizing organized artificial neural networks and inspired by the structure of the human brain. In terms of deep learning this structure called an artificial neural network (ANN). Since artificial neural networks are structured like a human brain consisting connected set of nodes similar to brain neurons, the human thinking way that can be learnt and mimicked by machines via a deep learning algorithm. Because of this hierarchical neural network structure of deep learning approach, the process of data can be carried out through a nonlinear approach. When some input data is fed to a neural network, process via layers of the perceptron to produce the desired output is done. The mathematical structure of the deep learning includes linear algebra, probability and information theory (stochastic models), and numerical computations. In mathematical structure of a neural network, data of one pattern set as the input set in entered to a set of nodes via some factors named weights which are selected randomly, then the output of these nodes which are called hidden layer is given to the activation function, and the output of these activation function box will be the network output. The architecture of a node in the neural network is represented as figure2 [8].

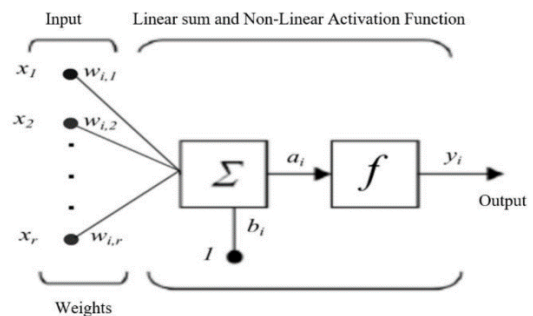


Fig. 2. Neural network Node Architecture

There are two steps for training in all models of deep learning and also in CNN, feed-forward step and back-propagation step. In deep learning, the learning procedure is done by updating the weight values. Backpropagation or “backward propagation of errors” is introduced for solving the problem of updating the weights and training multiple neural networks or deep neural networks [9]. In the backpropagation algorithm, the errors are passed to hidden layers to train the network. If the error does not reach the hidden layers, weight cannot be adjusted. So there would be no point if they were not trained. The training problem can be solved by the activation function. The calculation procedure in deep neural networks is as follow:

Step1- initialize the weight

Step2- calculate the error (can be specified as the mean squared error (MSE)), e stands for error.

$$e = \frac{1}{N} \times \sum_{i=1}^N [t_i - o_i]^2 \tag{1}$$

Step3- adjust the weight values to reduce the error, w stands for weight.

$$w(n) = -\eta \frac{\partial e}{\partial w} + \alpha w(n-1) \tag{2}$$

$$\Delta w = -\eta \frac{\partial e}{\partial w} \tag{3}$$

Error is the difference between the correct output and network output. The weight adjustment is done based on the error.

Convolution is a mathematical operation as follow:

$$x[n] * h[n] = \sum_{k=-\infty}^{+\infty} x[n-k].h[k] = \sum_{k=-\infty}^{+\infty} x[n].h[n-k] \tag{4}$$

$$(x * h)(t) = \int_{-\infty}^{+\infty} f(\tau)g(t - \tau)d\tau = \int_0^t f(\tau)g(t - \tau)d\tau \tag{5}$$

In a CNN, for an arbitrary input an arbitrary filter is used to perform this convolution operation. The performance of this filter is represented as figure3.

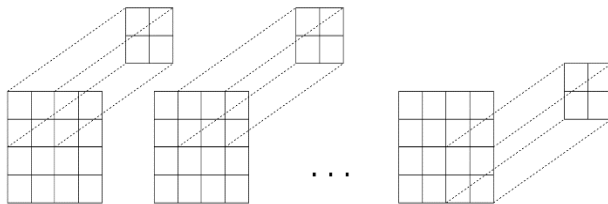


Fig.3. Convolution operation in CNN

During convolution performance, the input is scanned by filter according to its dimension. In other words, the filter is going to convolve across each filter. the CNN consists of three main layers: convolutional layer, pooling layer, and fully connected layer respectively. convolutional layer receives input then transforms the input in some way and then output the transformed input in the next layer. During the pooling layer, a function commonly based on the maximum operation is applied over a portion of the input to create an output. Out of pooling layer is smaller than the input dimensionally. Behind the convolution and pooling layer, there is another layer named fully connected layer which has a neural network construction and features and representations that are learned from data through Convolutional and pooling layer are fed to this fully connected neural network without hand engineering them like the event that is happened in machine learning, one or multiple convolutions and pooling layer can be in a CNN. In other

words, the output of the pooling layer is the input of the future artificial neural network. Structure of a CNN is represented as figure4.

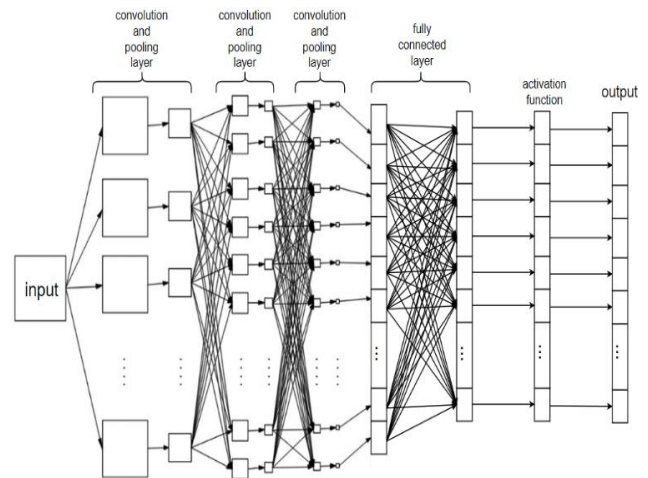


Fig.4. CNN architecture

In a CNN, the gradient descent is the mathematical facts that help to find the best filters, convolutions, pooling layers, fully connected layers etc. It starts with random values for everything such as filters, convolution layers, then it checks the lots of errors and then uses the gradient descent to move in a direction to reach a few errors. And this gradient descent is used in the back-propagation stage.

3.2. FLOW CHART

Deep neural network is an extension of neural network that contains two or more networks. The model of the deep neural network is represented as figure5.

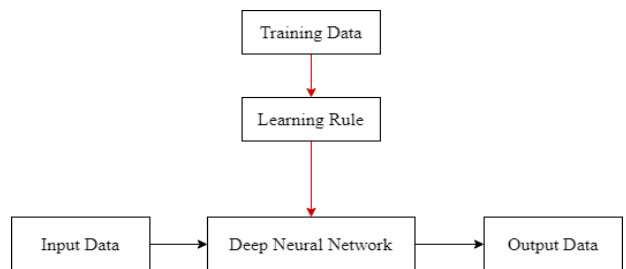


Fig.5. Deep neural network training model

The training data train the network using learning rules. Once the network is trained, the input data is provided and the network generates output data. In neural networks all data is stored in weights, so for training the network with new information, the weight values should be modified. The systematic way of modifying the weight values is called "Learning Rule". Descent gradient is utilizing to update the weight values in back-propagation stage. Back-propagation stage is executed to reduce the error value and reach the best result as output in Deep learning. The use of back-propagation

as the crucial part of deep learning algorithm is represented as figure 6.

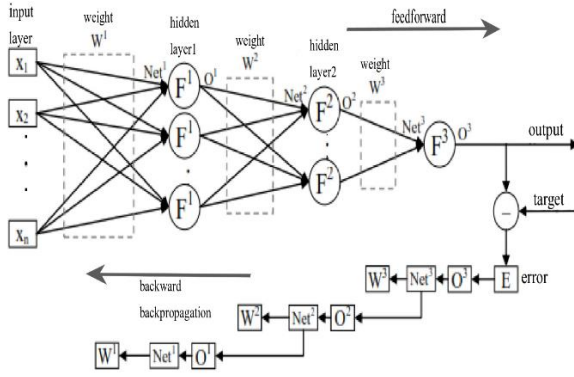


Fig.6. BP process

In CNN which consists three main layers, first of all, the input is convolved by the filters, then it crosses the pooling layer to be smaller than input in dimension, convolution and pooling layers are three-dimensional layers and one CNN may have multiple convolutions and pooling layers. Then the output of the last pooling layer is flattened. The output of pooling layer is 3D feature map, but the input of fully connected layer is 1D feature vector. Flattening turns the 3D output of pooling layer into 1D input of fully connected layer. After this process, the learning process is completed and learned data is entered to a neural network and finally, the output is obtained as it is shown in figure4. The flowchart of CNN is represented as figure7.

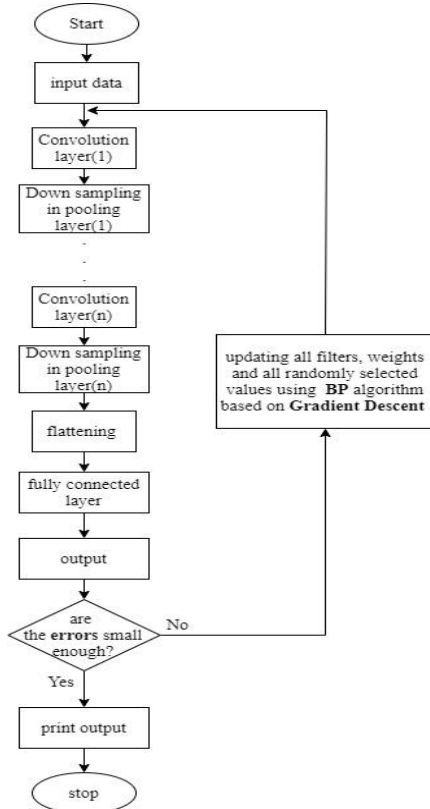


Fig.7. CNN flowchart

4. COMPARISONS ON THE AI-ALGORITHM USING DLA

The meaning of AI is to enable the machine to think and act like a human and that is the main target of as an AI system like self-driving cars, robots etc. machine learning is a subset of AI systems that makes that system to be AI. And it provides statistical tools to explore and analyze the data. Machine learning is divided into three categories: supervised, unsupervised, and reinforcement. The role of deep learning as a subset of machine learning is determining the learning procedure by mimicking the human brain. The main idea behind deep learning is imitating the human brain in the learning process and has an architecture like the human brain that is called multi neural network architecture. Deep learning has various techniques inside of itself. Finally, being AI is considered as a goal that deep learning helps to reach this aim.

5. CONCLUDING REMARKS AND DISCUSSIONS

According to above discussion, it can be stated that an artificial neural network is a functional unit of deep learning; deep learnings uses artificial neural networks which mimic the behaviour of the human brain to solve complex data-driven problems. Now deep learning in itself is a part of machine learning which folds under the larger umbrella of artificial intelligence. Artificial intelligence, machine learning, and deep learning are interconnected field. Machine learning and deep learning aids artificial intelligence by providing a set of algorithms and neural networks to solve data-driven problems. Deep learning makes use of artificial neural networks that behave similarly to the neural networks in our brain. Since deep learning has the ability to analyze big data and modelling the human brain using neural networks to more effectively realize how the human brain works, it has a specific position in cognitive systems and provides cognition for a system.

ACKNOWLEDGEMENT

Thanks to members of Istanbul Technical University, Cognitive System Laboratory.

REFERENCES

- [1] Pierson, P.E, Lillian. Data Science For Dummies, John Wiley & Sons, Inc, 2015.
- [2] *Relationship of artificial intelligence, machine learning and deep learning: 101.* (n.d.). Academia.edu - Share research. https://www.academia.edu/39460436/Relationship_of_Artificial_Intelligence_Machine_Learning_and_Deep_Learning_101
- [3] D. H. Tolley, (Mapelton, UT, US) 2007 Method of operating a plant and a control system for controlling the same United States GBI Structures, L.L.C. (Springville, UT, US) 7206772.
- [4] H.D. Wehle, (2017). Machine Learning, Deep Learning, and AI: What's the Difference?.
- [5] H. D. Wehle, (2019). Artificial Intelligence. 10.13140/RG.2.2.28657.74089.
- [6] T. Mitchell, (1997). Machine Learning. McGraw Hill. p. 2. ISBN 978-0-07-042807-2.

- [7] *CIS 419/519 applied machine learning - Spring 2020*. (n.d.). Penn Engineering | Inventing the Future. <https://www.seas.upenn.edu/~cis519>
- [8] B. Alafi, "A New Cognitive System Modeling with the Combination of Expert and Neural Network Systems". *The Journal of Cognitive Systems*, Vol.3, No.2, 2018, pp.34-35.
- [9] Goodfellow-et-al-2016, Deep Learning, Ian Goodfellow and Yoshua Bengio and Aaron Courville, MIT Press, url{<http://www.deeplearningbook.org>, 2016
- [10] H. Schulz, S. Behnke, "Deep Learning". *KI-Künstliche Intelligenz*, Vol.26, No.4, 2012, pp.357-363.
- [11] A. Mosavi, S. F. Ardabili, A.R. Varkonyi-Koczy, "List of Deep Learning Models," 2019, (doi: 10.20944/preprints201908.0152.v1)

BIOGRAPHIES

Behnaz Alafi obtained her BSc. degree in electrical engineering from Iran University of Science and Technology(IUST) in 2009, Tehran, Iran. MSc. Student at Istanbul Technical University. She has experience in electrical systems of oil and gas, petrochemical, and energy industry since 2009.