

What is Artificial Intelligence? Technical Considerations and Future Perception

 *Mustafa Ergen*

Department of Electrical and Electronics Engineering, Istanbul Technical University, Istanbul-Turkey

Keywords: AI, deep learning, diagnostics, machine learning, robotics

Introduction to artificial intelligence

Today, Artificial Intelligence (AI) is another, however, strong technological wave that is flattening the world by providing the ability for a machine to perform cognitive functions, such as perceiving, reasoning, learning and interacting. AI has rapidly entered our lives by solving business problems due to three technological developments that have reached enough maturity and convergence: (1) advancement in algorithms, (2) massive data, and (3) increasing computational power and storage at low cost. In the eighties, a similar wave transformed the world with the personal computer technologies, where computational power became very cheap and affordable. In the same way, AI makes prediction cheap and affordable and will lead to immediately automating the routine and reproducible works through machines.

Although the forecasts are updated every day, artificial intelligence is expected to increase the average annual revenues by more than ten trillion dollars all over the world and expected to transform many industries. In order to predict how wide the effect will be, we must first look at what the is difference with respect to non-artificial intelligence, real intelligence (1).

Intelligence, as the most distinctive feature of human beings, is the one unique feature that separates us from other creatures. Today's question is to what extent can we realize intelligence in an artificial environment!

Human intelligence works in a certain hierarchy. This layered structure allows us to understand patterns. It has the ability to match them to a symbol and makes logical inferences with each symbol. We call these patterns "ideas" and the "knowledge" is when these ideas connect with each other. Being able to speak, to symbolize the words, and to transfer the information to another emerge as the most important intelligent features of a human being. Beyond that, intelligence is the ability to make logical

inferences, solve problems, adapt, interact, and learn. This is accomplished by means of mental functions, such as perception, memory utilization, attention, speech, and planning. Judgments for instance utilize intuition and empathy which are also considered moral norms. People learn from their past experiences, adapt to new situations, advance with creativity, discuss abstract ideas, and use their knowledge to change their surroundings (1).

Today, AI focuses on a specific functionality of intelligence-finding patterns. As in intelligence, an algorithm that mimics the brain is required. Furthermore, in order to find patterns—we must teach the patterns first. Hence, data is necessary. Further, as in the brain, we need to make computations.

Thus far, the AI wave has been held back due to the lack of data as well as the costly computation power. The information required for machines to learn is accumulated by the personal computer and Internet phase, and today people and machines together started to double the data every year. And, the capacity of computational power is becoming compatible with that of the human brain. For example, the human brain is one-sixth of a basketball and processes 1 million trillion transactions per second. Computers that are the size of a basketball court can handle 93,000 trillion transactions per second. One can conclude that a single task can be handled by machines and, according to Moore's Law, as computers double their processing power every 18 months, the trend will reach the capacity of the human brain in 2025 (2).

What remains is the sufficiency of the algorithms. Although the algorithms are based on the statistical methods that date back to the eighteen hundreds, which used to determine the orbits of the planets, the perceptron put forward by Professor Frank Rosenblatt in 1956 became the basis of today's artificial intelligence. The decision-making process is mathematically con-

Address for correspondence: Mustafa Ergen, Department of Electrical and Electronics Engineering, Istanbul Technical University, Istanbul-Turkey

Phone: +90 554 746 28 13 E-mail: mustafaergen@itu.edu.tr

©Copyright 2019 by Turkish Society of Cardiology - Available online at www.anatoljcardiol.com
DOI:10.14744/AnatolJCardiol.2019.79091



structured with a mechanism that makes the decision of yes or no, by multiplying the incoming data with a weight vector. And, this weight vector is adjusted during the training phase with known labeled datasets. Over the years, AI algorithms have advanced with major breakthroughs.

History of artificial intelligence

For the past century, the AI wave has reached glorious peaks and silent downs. The hype led agencies and governments to spend money on artificial intelligence development, however, unsatisfactory results shut down these initiatives. Maturity of the algorithms needed to wait for data and data needed to wait for prevalent computational power (3).

In the 1920s, AI was just science fiction. By the 1950s, a small step was made, moving AI from hype to reality. Alan Turing published a paper titled "Computing Machinery and Intelligence" and introduced Imitation Game (aka Turing Test) for determining whether a machine is intelligent or not. The test is very intuitive that if a human judge when engaged in an interaction with a computer cannot reliably distinguish machine from human, the machine passes the test.

The term "Artificial Intelligence", as a respectable scientific term, was coined in the "Dartmouth Conference" of 1956, organized by Marvin Minsky, John McCarthy, and Shannon and Nathan Rochester. One year later, after introduction of perceptron, American computer scientist Arthur Samuel coined the term "machine learning".

Introduction of deep learning came with Ukrainian mathematician Alexey Grigorevich in 1965, by stacking several perceptrons on top of one another—resembling today's deep-learning architectures inspired from Neural Networks (NN) of the human brain.

There were a few notable successes before AI went into hibernation in the 1970s, which continued for a decade.

By 1980's, Geoffrey Hinton et. al. introduced backpropagation, a mechanism to self-optimize Artificial NN without human intervention. His success was adjusting the weights of the NN nodes across multiple layers.

In 1989, Yann LeCun introduced the Convolutional Neural Network (CNN) for image recognition. In 1992, Bernhard E. Boser, Isabelle M. Guyon, and Vladimir N. Vapnik introduced Support Vector Machines (SVM), so that classifying the sentiment and understanding of human speech became easy for natural language processing.

By 1991, Sepp Hochreiter introduced a Recurrent Neural Network (RNN) to be useful in sequencing tasks. And in 1997, Hochreiter and Jurgen Schmidhuber introduced Long Short-Term Memory (LSTM) for advanced speech to text translation.

In 2006, Hinton pre-trained the network with Deep-belief Network before employing backpropagation and coined the phrase "deep learning".

By 2010, AI propelled with an increasing boom of AI products and mass adoption began especially with the ImageNet⁽¹⁾ competition of 2012, won by Hinton's team by training a CNN with 1.2 million images on two GPU cards and with DeepMind's Reinforcement Learning Algorithm over CNN-first trained for Atari games to play at human level in 2013 then in 2016, under the umbrella of Google, beats GO world champion⁽²⁾ (4).

Today AI is machine learning!

The vast majority of the AI advancements and applications are categorized today as machine learning (ML). ML algorithms are statistical instruments that are used to find patterns in massive amounts of data and then these patterns are used to make predictions. Patterns are relationships between the input and the output (5).

For instance, in order to recognize objects, ML can be used as a classifier. First, a training procedure is built by a sample data so that ML algorithms are tuned with the right parameters. Training is established with known data and its associated labels. Known data is first processed for feature extraction and then fed into the ML algorithm. The ML algorithm will find the pattern and classify it. This is called supervised learning. For new images fed into the model, the machine will classify accordingly. There is also unsupervised learning (clustering), where learning seeks patterns in the data that have no labels and reinforcement learning (robot planning), where learning is with trial and error (6).

Deep learning is state of art machine learning!

Deep Learning (DL) is a subset of machine learning and designed to mimic the network of neurons in a brain. Artificial Neural Networking (ANN) is an architecture where there is a hidden layer between input and output and the deep neural network (aka deep learning) utilizes more than one hidden layer, where the layers are stacked on top of each other. A layer is composed of nodes. Each node is connected to nodes in the previous and next layers as in brain. Each input is multiplied by a weight in each neuron. Multiplied outputs are input for the next layer. The

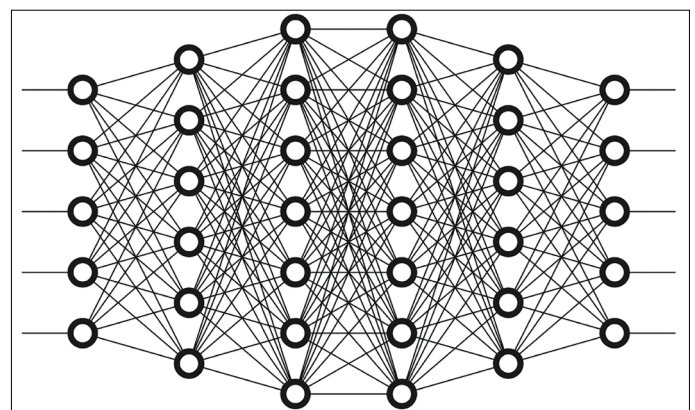


Figure 1. Schematic representation of neural network

1. <http://image-net.org/challenges/LSVRC/>

2. <https://deepmind.com/research/case-studies/alphago-the-story-so-far>

final layer is the output layer and provides the probability of each class for the classification task. These multiple hidden layers help optimization escape from bad local optimums.

ANN is inspired by the human brain with the training of more than one hidden layer, introduced in 2006 by Geoffrey Hinton for unsupervised training (6, 7). Before this, technically, adjusting the weights by backpropagation was an issue (8).

The training procedure is for updating the weight of all the nodes. The difference between ML and DL is in the feature extraction. Typically, an image, for instance, can have many features and not all are relevant to the algorithm. In ML, there is a process, before feeding into the algorithm, to find a relevant set so that the system learns something particular and meaningful. This is called feature extraction and the main purpose is to reduce the dimension of the problem. Principal Component Analysis is a well-known technique, where the dimension is reduced with respect to eigenvalues and the histogram of oriented gradients (HOG) is a featured descriptor used in image processing (8).

DL eliminates the CNN procedure, in which the first layer learns small details and the next layers use the previous information to construct complex information. CNN utilizes filters for feature extraction. RNN further focuses on temporal data sequences and again uses the multilayered neural network, assuming the inputs are correlated.

With ML, you need fewer data to train the algorithm than DL. DL requires an extensive and diverse set of data to identify the underlying structure. Besides, ML provides a faster-trained model. Most advanced DL architectures can take days to a week to train. The advantage of DL over ML is its high accuracy. You do not need to understand which features are the best representatives of the data since the neural network knows how to select critical features. In ML, you need to choose for yourself which features to include in the model. In order to overcome the drawback of insufficient data and reduce the training time, Transfer Learning (TL) is gaining popularity. TL is the re-use of knowledge from one problem and applying it to a different but related problem. TL uses the pre-trained model for another problem. For instance, if there is a network that wants to learn to identify dogs, you can use an existing network that is commonly used for image recognition of other animals, and fine-tune it to train the network with dogs (9, 10).

Conclusion - Why AI is important?

AI is improving humanity in the fields of many areas, including resource efficiency, autonomous machines, healthcare, agricultural etc. By performing simple fixed tasks, AI is replacing human workers for certain jobs as the single most powerful force of our time. This not only increases productivity but is also transforming the jobs. Expectations in AI will create more jobs than it will destroy (11, 12).

The AI hype will die eventually when AI becomes another thread as in the Internet, electricity, and transportation. Today, AI is promising, or in some cases threatens, changes in busi-

ness and society at large. The industrial era let machines do the physical work; the information era enabled machines to do the computation and storage, now the AI era will let machines make the decisions (13).

Today, we are also seeing that some AIs are already passing the Turing Test (14)!

Conflict of interest: None declared.

References

1. Ergen M. Yapay zeka ve Kristof Kolomb. Cited 2019 Oct 9. Available from: <https://m.platinonline.com/yazarlar/mustafa-ergen/yapay-zeka-ve-kristof-kolomb-943638>.
2. Cited Oct 2019. Available from: <https://www.technologyreview.com/s/612768/we-analyzed-16625-papers-to-figure-out-where-ai-is-headed-next/>.
3. Anyoha R. The History of Artificial Intelligence. Cited 2019 Aug 28. <http://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/>
4. Cited 2019 Oct 9. Available from: <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/an-executives-guide-to-ai>.
5. Cited 2019 Oct 9. Available from: <https://sonix.ai/articles/difference-between-artificial-intelligence-machine-learning-and-natural-language-processing>.
6. Retrieved from <https://aircto.com/blog/everything-you-need-to-know-about-artificial-neural-network-deep-learning-part-i/>.
7. Brownlee J. How to Fix the Vanishing Gradients Problem Using the ReLU. Cited 2019 Jan 11. Available from: <https://machinelearningmastery.com/how-to-fix-vanishing-gradients-using-the-rectified-linear-activation-function/>.
8. How does deep learning work and how is it different from normal neural networks applied with SVM? How does one go about starting to understand them (papers/blogs/articles)? Cited 2019 Sep 28 2017. Available from: <https://www.quora.com/How-does-deep-learning-work-and-how-is-it-different-from-normal-neural-networks-applied-with-SVM-How-does-one-go-about-starting-to-understand-them-papers-blogs-articles>.
9. Transfer Learning - Machine Learning's Next Frontier. Cited 2017 March 9. Available from: <http://ruder.io/transfer-learning/>.
10. Martin S. What Is Transfer Learning? Cited 2019 Feb 7. Available from: <https://blogs.nvidia.com/blog/2019/02/07/what-is-transfer-learning/>
11. Cited 2019 Oct 9. Available from: <https://www.forbes.com/sites/bernardmarr/2018/07/27>.
12. Costello K. I&O leaders need to strategically leverage AI as a core accelerant to digital business initiatives. Cited 2019 feb 13. Available from: <https://www.gartner.com/smarterwithgartner/gartner-predicts-the-future-of-ai-technologies/>.
13. Hao K. What is AI? We drew you a flowchart to work it out. Cited 2019 Nov 10. Available from: <https://www.technologyreview.com/s/612404/is-this-ai-we-drew-you-a-flowchart-to-work-it-out/>.
14. Nieva R. Alphabet chairman says Google Duplex passes Turing test in one specific way. Cited 2019 May 10. Available from: <https://www.cnet.com/news/alphabet-chairman-says-google-duplex-passes-turing-test-in-one-specific-way-io-2018/>.