Saanvi Bhargava

Dr. Nelson

ATCS: Neural Networks P2

9 February 2024

Minds and Machines: Tracing the Evolution of Artificial Neural Networks

As society stands at the crossroads of innovation, tracing the path from the creation of Rosenblatt's Perceptron and Parallel Distributed Processing (PDP) to the ever-evolving landscape of current AI, the journey into the origin of Artificial Intelligence becomes truly riveting. The human fascination with understanding the brain and intelligence has consistently inspired endeavors to emulate these phenomena. In the past, scientists have embarked on a path that aimed to model the workings of the human brain. Consequently, a substantial portion of the following research and achievements in this domain emerged from individuals with expertise in psychology, neuroscience, and computer science. The first major discovery relating to neural networks occurred in the early 1940s, with McCulloch and Pitts.

In 1943, neurophysiologist Warren McCulloch and mathematician Walter Pitts published a groundbreaking article. This work, notable for its time, discussed a theory of the function of the human nervous system, establishing the first ever mathematical model of a biological neuron (McCulloch and Pitts 100). This paper was so piquant to creative scientists and psychologists of that era since it was the first tangible computational analogy of the brain.

McCulloch and Pitts used four fundamental principles of neuroscience as the basis of their model: the brain is composed of neurons, neurons connect through junctions called synapses, neurons can transmit signals to one another, which can be either stimulating or inhibitory, and a neuron only fires once it reaches a threshold of stimulation (Dukes, “McCulloch-Pitts”). Their simple mathematical model and these four principles were considered the beginning of artificial neural network history.

Curiously, just a few years prior, in 1936, Alan Turing introduced the concept of the Turing Machine as a framework for universal computation, and in 1950, Turing created the Turing Test for detecting Artificial Intelligence. These few events already reveal a pattern of the early to mid-twentieth century: a clear probing into the relationship between the brain, mathematics, and reasoning.

In 1949, psychologist Donald Hebb proposed the principle, now known as Hebbian learning, that when two neurons fire together, their connection strengthens in the brain and in a potential mathematical model. Donald Hebb's groundbreaking work laid the foundation for understanding synaptic plasticity, the ability of the connections between neurons (synapses) to change in strength (Brown and Milner). Hebb proposed this principle in his 1949 book, "The Organization of Behavior." This concept has since become fundamental in neuroscience and was a significant source of inspiration and ideation for Rosenblatt.

Frank Rosenblatt, a psychologist and computer scientist, was born in 1928 in a suburb of New York City. After studying at the Bronx High School of Science, he attended Cornell, where he majored in social psychology and earned his Ph.D. in psychology. In the late 1950s, Rosenblatt laid the foundation for neural networks with his pioneering work in creating the Perceptron (Zhang 5). His interdisciplinary educational background enabled him to make revolutionary connections between neuroscience and computing.

The 1959 Mark I perceptron, Rosenblatt's brainchild, aimed to simulate the functionality of a single neuron in the brain through a supervised learning algorithm. The Mark I Perceptron is currently housed in the Smithsonian National Museum of American History (Lefkowitz). The three-layered perceptron network contains an array of 400 photocells, resistors that change resistance depending on the amount of light on it. Operating as a single-hidden-layer neural network, the perceptron utilizes an algorithm designed for the classification of specific inputs into two categories (Dukes, “Rosenblatt’s Perceptron”). It accomplishes this task through a large number of training iterations, allowing it to make predictions based on learned patterns. Rosenblatt originally created a feedforward network with multiple characteristics, the most basic of which is that it requires a perceptron of an input layer, at least one hidden layer, and an output layer (“Feed-forward networks”). Each neuron in a layer is linked to every neuron in the following layer, creating the continuous “feed-forward” flow of information (Hardesty). There must be no connections between neurons in the same layer.

McCulloch Pitts Neuron
(assuming no inhibitory inputs)

$$y = 1 \quad if \sum_{i=0}^{n} x_i \geq 0$$
$$= 0 \quad if \sum_{i=0}^{n} x_i < 0$$

Perceptron

$$y = 1 \quad if \sum_{i=0}^{n} w_i * x_i \geq 0$$
$$= 0 \quad if \sum_{i=0}^{n} w_i * x_i < 0$$

Figure 1: McCulloch Pitts vs. Rosenblatt Models (Chandra)

Compared to McCulloch and Pitts' model from years prior, The Mark I perceptron incorporated activation weights, bias, learning, and numerical inputs. Rosenblatt's perceptron accounted for not all inputs having the same strength, determined by the activation weights and bias, depicted in Figure 1. Activation weights are weight values in an activation function for a neuron that represent the strength and direction of the connections between the input and output of the neuron. The MCP model also only accommodated binary inputs, and the Mark I perceptron was able to take numbers as inputs. This advancement meant the range of tasks the model could complete was expanded exponentially. The ability of Rosenblatt's model to train and learn to adjust weights is unique from the MCP model. From Hebb's insights, Rosenblatt deduced that adjusting the perceptron's weights would be necessary based on the difference between the desired and actual output. His multi-layer perceptron (MLP), originating from an evolving understanding of the human brain, allowed the neural network to grasp and learn linear connections within the provided data. Researchers aimed to create models that closely mirrored the intricate structure of the neurons in the brain. By incorporating many layers of interconnected neurons, the goal was to capture and emulate the complexity of neural processes. The introduction of hidden layers in the MLP enabled neural networks to handle intricate patterns and relationships within data better, ultimately enhancing their capacity to tackle slightly more complex tasks ("CHATGPT, AI, and implications for higher education"). Although its ability to learn behavior and classify input was groundbreaking, this invention had significant limitations in its ability–the perceptron could only solve simple, linear tasks.

In response to these challenges faced by Rosenblatt's perceptron, scientists hoped to improve its behavioral abilities. The proposed solution came in the form of the backpropagation algorithm, now the most widely used tool in the field of neural networks. In a perceptron, when training the model, it is necessary to calculate the error considering the weights, input, and expected and actual output. Backpropagation is the algorithm for calculating the gradients of the cost function with respect to the weights. It is required in non-linear MLPs because there may be multiple minimums for the loss function. Backpropagation is used to improve the output of non-linear MLPs by propagating the error in a backward direction and calculating the gradient of the cost function for each weight. These gradients are used in the process of gradient descent. Gradient descent is an optimization technique employed to identify the weights associated with a cost function ("Cauchy and the Gradient Method"). The objective is to iteratively move down the cost function until its lowest point is reached, revealing the optimal weights. This process relies on the gradient and the learning rate to navigate the cost function. The gradient guides the algorithm towards the direction leading to the minimum point of the cost function, while the learning rate dictates the pace of this descent. Once the minimum point is reached, gradient descent identifies the weights at that optimal position.

Backpropagation was first introduced in the 1960s and popularized decades later by Rumelhart, Hinton, and Williams. David Everett Rumelhart was a pioneer in the field of cognitive neuroscience and explored the idea that no single neuron in the human brain does its job alone in processing information (Carey). He created computer models in the 1970s and 1980s that simulated human perception, language understanding, memory and a wide range of other cognitive tasks. Rumelhart and his colleague, David Clarence McClelland, co-authored a book titled "Parallel Distributed Processing: Explorations in the Microstructure of Cognition." This work significantly increased the exposure of interconnected neural networks to a broader audience, including psychologists, neuroscientists, and computer scientists. He became interested in the principle that human thoughts emerge from activity in and interaction between neurons. Similarly, Paul John Werbos proposed the advancement of reinforcement learning systems that are both more potent and biologically credible. He introduced the novel concept of employing neural networks to approximate dynamic programming (ADP), encompassing a range of optimization control methods (Werbos 1550). In order to implement ADP, he created an algorithm later called backpropagation. Like other psychologists and computer science, he aims to develop designs to explain the intelligence in the human brain and in human experiences ("Paul Werbos Biography").

Even with the number of advancements made from 1943 to the 1970s, the landscape of artificial intelligence encountered a significant setback. The first AI winter was from 1974 to 1980 and was a result of the limited capabilities of AI programs at the time. Computing power could only handle trivial versions of the more difficult problems they were tasked with. Marvin Lee Minsky, an American cognitive and computer scientist concerned largely about the research of artificial intelligence, wrote a paper suggesting that there could not be an extension from the single-layered neural network to a multi-layered neural network (Thorwirth). This negative take on Artificial Intelligence contributed significantly to the AI winter, and research efforts and funding went down drastically.

The past trajectory of artificial neural networks, from McCulloch and Pitts to 21st-century advancements, is a journey of exploration and discovery, with milestones in fields spanning psychology, computer science, and neuroscience. In the 1940s, Warren McCulloch and Walter Pitts created the first mathematical model of a biological neuron, inspiring subsequent research in neural networks. Donald Hebb's principle of synaptic plasticity furthered the understanding of neural connections. Frank Rosenblatt's Perceptron, developed in the late 1950s, marked a significant milestone in AI history, laying the groundwork for neural networks. However, Rosenblatt's perceptron had limitations in solving complex tasks, which led to the creation of backpropagation. The introduction of the backpropagation algorithm, pioneered by Rumelhart, Hinton, and Williams, optimized weights based on error calculations. Despite these advancements, the first AI winter occurred from 1974 to 1980, caused by limited computing power and skepticism about the potential of multi-layered neural networks, specifically by Marvin Minsky. Nevertheless, researchers persisted in exploring neural networks and reinforcement learning, leading to subsequent AI breakthroughs. Looking ahead, the future of neural networks holds no bounds. One thing remains certain: it remains essential to draw inspiration from the inventors of the past and push the boundaries of what is possible.

Bibliography

Brown, Richard E., and Peter M. Milner. “The legacy of Donald O. Hebb: More than the hebb synapse.” *Nature Reviews Neuroscience*, vol. 4, no. 12, Dec. 2003, pp. 1013–1019, https://doi.org/10.1038/nrn1257.

Carey, Benedict. “David Rumelhart Dies at 68; Created Computer Simulations of Perception.” *The New York Times*, The New York Times, 19 Mar. 2011, www.nytimes.com/2011/03/19/health/19rumelhart.html.

Chandra, Akshay L. “Perceptron: The Artificial Neuron (an Essential Upgrade to the McCulloch-Pitts Neuron).” *Medium*, Towards Data Science, 11 Aug. 2018, towardsdatascience.com/perceptron-the-artificial-neuron-4d8c70d5cc8d.

“CHATGPT, AI, and Implications for Higher Education: History of AI and Neural Networks.” *History of AI and Neural Networks - ChatGPT, AI, and Implications for Higher Education*, libguides.aurora.edu/ChatGPT/History-of-AI-and-Neural-Networks. Accessed 8 Feb. 2024.

Dukes, Kais. “McCulloch-Pitts: The First Computational Neuron.” *LinkedIn*, 25 June 2023, www.linkedin.com/pulse/mcculloch-pitts-first-computational-neuron-dr-kais-dukes?trk=article-ssr-frontend-pulse_little-text-block.

Dukes, Kais. “The History of Neural Networks Part 2: Rosenblatt’s Perceptron.” *LinkedIn*, 28 June 2023, www.linkedin.com/pulse/history-neural-networks-part-2-rosenblatts-perceptron-dr-kais-dukes#:~:text=Frank%20Rosenblatt,computers%20to%20learn%20from%20experience.

“Feed-Forward Networks.” *Neural Networks - Architecture*, Stanford CS, cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/Architecture/feedforward.html. Accessed 8 Feb. 2024.

Hardesty, Larry. “Explained: Neural Networks.” *MIT News*, MIT News Office, 14 Apr. 2017, news.mit.edu/2017/explained-neural-networks-deep-learning-0414.

Lefkowitz, Melanie. "Professor's Perceptron Paved the Way for AI – 60 Years Too Soon." *Cornell Chronicle*, 25 Sept. 2019, news.cornell.edu/stories/2019/09/professors-perceptron-paved-way-ai-60-years-too-soon.

McCulloch, Warren S., and Walter Pitts. "A logical calculus of the ideas immanent in nervous activity." *Bulletin of Mathematical Biology*, vol. 52, no. 1–2, Jan. 1990, pp. 99–115, https://doi.org/10.1007/bf02459570.

"Paul Werbos Biography." *NSS*, nss.org/paul-werbos-biography/. Accessed 8 Feb. 2024.

Werbos, Paul J. "Backpropagation through time: What it does and how to do it." *Proceedings of the IEEE*, vol. 78, no. 10, Oct. 1990, pp. 1550–1560, https://doi.org/10.1109/5.58337.

Zhang, Youheng. "A historical interaction between artificial intelligence and philosophy." *Teorie Vědy / Theory of Science*, vol. 1, no. 1, 23 Oct. 2023, https://doi.org/10.46938/tv.2023.579.