Aeliya Grover

Dr. Nelson

ATCS: Neural Network, Period 4

September 14 2023

Exploring The Formation and Development of Rosenblatt's Mutli-layer Perceptron

The development of Rosenblatt's Multi-layer Perceptron can best be explained through the inspiration from biological neurons and the goals of model improvement. As knowledge of the neuron improved, so did the neural network models. In the early 1900s, scientists knew that the neuron consisted of four main parts: the dendrites, the soma, the axon, and the synapse. Working together, "a neuron takes an input signal (dendrite), processes it like the CPU (soma), passes the output through a cable like structure to other connected neurons (axon to synapse to other neuron's dendrite)...This massively parallel network also ensures that there is a division of work. Each neuron only fires when its intended criteria are met" (Chandra). The dendrites of the neuron cell take inhibitory and excitatory signals. With these signals, the soma determines if an all-or-nothing signal should be sent down the axon to the synapse to release neurotransmitters for another neuron. Inhibitory signals push to prevent a signal from going, while an excitatory signal pushes to release the signal (Cherry).

In 1943, Warren McCulloch, a neuroscientist, and Walter Pitts, a logician, decided to make the first mathematical model of a neuron, creating the first artificial neuron. The artificial neuron processes the inputs by taking in a binary input, either '0' representing an inhibitory signal or '1' representing an excitatory signal. The output value is calculated by an activation function and is sent as an output (Chandra). Some limitations of this model include the inability to prioritize some inputs over others, accepting only binary inputs and outputs, and having no form of learning or adaptation. However, the McCulloch-Pitts neuron model set the groundwork for artificial neural network connectivity. The mathematical model and information processing system were heavily based on the similar functions of the dendrites, soma, axon, and synapsis of the biological neuron, creating the starting point of what would later become Rosenblatt's Multi-layer Perceptron.

New findings of Hebbian theory and neuron capabilities later influenced artificial neural networks (ANN), as seen through Rosenblatt's 1957 perceptron. Donald Hebb proposed that when biological neurons fire at one another, they build extra myelin sheets along the neuron's axon, allowing neurons to increase the speed of their signal. This then improves the signaling process of the two neurons, allowing more weight, or importance, to a specific neuron signal and pathway (Keysers and Gazzola). Using the newfound knowledge, Rosenblatt helped improve the development of neural networks through his work on perceptrons. Like weights in Hebbian theory and biological neurons, instead of sending the inputs and having them each contribute equally, weights and biases would be applied to the inputs to influence the effect of input connections. In his first model, there was only one input layer for inputs into the system, one hidden layer for calculations, and one output layer for a result. A hidden layer is located between the input and output layers of the algorithm, giving an output that goes directly to another layer (Hardesty). Improving from the McCulloch-Pitts neuron model, the perceptron included a simple error correction for weight and bias adjustment by comparing the model outputs to expected outputs, allowing error minimization (Schuchmann). With further influences from the biological neuron and advancements through overcoming limitations in previous artificial neuron models, a step was taken in developing Rosenblatt's Multi-layer Perceptron.

However, as limitations in Rosenblatt's Perceptron arose, doubts about its abilities led to the start of the first artificial intelligence (AI) winter as “in 1969, Minsky pointed out the limited representational abilities of a perceptron (to be exact, a single-layer perceptron cannot implement the classic XOR logical function), and despite not being a general critique about neural networks, this also contributed to global funding cuts in neural networks research” (Toosi *et al.*).

Because of exaggerations and overconfidence in the model’s capabilities, the eventual underperformance of scaling up to the XOR model led to a sharp decline in AI popularity. The decline in popularity then turned into budget cuts, and without the funding, development and research stagnated.

The key finding that allowed the exit from the first 1970s AI winter was the creation of Paul Werbos’ backpropagation. In his paper, Paul Werbos explains the process of backpropagation:

> In basic backpropagation, we start with arbitrary values for the weights. …[N]ext, we calculate the outputs… and the errors…for that set of weights. Then[,] we calculate the derivatives of [the errors] with respect to all the weights...If increasing a given weight would lead to more error, we adjust that weight downwards. If increasing a weight leads to less error, we adjust it upwards. After adjusting all the weights up or down, we start all over, and keep on going through this process until the weights and the error settle down...[T]he uniqueness of backpropagation lies in the method used to calculate the derivatives exactly for all of the weights in only one pass through the system. (Werbos)

Backpropagation overcame previous limitations of non-linearly separable equations, allowing for deeper training and additional layers. While Werbos had initially discovered backpropagation in the early 1970s, it was not well known until the 1980s.

Using backpropagation, David Rumelhart, George Hinton, and James McClelland used parallel distributed processing (PDP) in the 1980s to further help end the AI winter. Their neural networks focused on parallel processing, implementing backpropagation, and adding many more nodes and layers (McClelland *et al.*). This improved the multi-layer perceptron as more layers could be trained efficiently (Schumann). By incorporating backpropagation and a more complex error correction, the multi-layer perceptron overcame the initial problem of being unable to solve non-linear separable equations. Like the previous networks, the model is "organized into layers of nodes, and they're 'feed-forward,' meaning that data moves through them in only one direction. An individual node might be connected to several nodes in the layer beneath it, from which it receives data, and several nodes in the layer above it, to which it sends data" (Hardesty).

The scientific community adopted PDP models to model human cognition and processes such as memory, learning, and perceptions (Hardesty). The multi-layer perceptron neural network can be seen almost everywhere in modern-day applications (Schuchmann). Rosenblatt's Multi-layer Perceptron required many models and challenges to overcome. Starting with a simple neuron representation with a binary input and output configuration, weights prioritizing specific inputs were added, and further, more advanced error adjustment through backpropagation allowed for the creation of Rosenblatt's Multi-layer Perceptron.

Works Cited

Chandra, Akshay L. "McCulloch-Pitts Neuron - Mankind's First Mathematical Model Of A Biological Neuron." *Towards Data Science*, 24 July 2018, towardsdatascience.com/mcculloch-pitts-model-5fdf65ac5dd1. Accessed 12 Sept. 2023.

Cherry, Kendra. "Action Potential and How Neurons Fire." *verywellmind*, www.verywellmind.com/what-is-an-action-potential-2794811. Accessed 12 Sept. 2023.

Hardesty, Larry. "Explained: Neural networks." *MIT News*, 14 Apr. 2017, news.mit.edu/2017/explained-neural-networks-deep-learning-0414#:~:text=Neural%20networks%20were%20first%20proposed,the%20first%20cognitive%20science%20department. Accessed 12 Sept. 2023.

Jia, Allison. "History of Parallel Distributed Processing." *Harker Horizon*, no. 3, 2021, horizon.harker.org/perspectives/2020/10/28/History-of-Parallel-Distributed-Processing.html. Accessed 12 Sept. 2023.

Keysers, Christian, and Valeria Gazzola. "Hebbian learning and predictive mirror neurons for actions, sensations and emotions." *PubMed Central*, www.ncbi.nlm.nih.gov/pmc/articles/PMC4006178/#:~:text=The%20term%20Hebbian%20learning%20derives,growth%20process%20or%20metabolic%20change. Accessed 12 Sept. 2023.

Laszlo, Sarah, and David C. Plaut. "A Neurally Plausible Parallel Distributed Processing Model of Event-Related Potential Word Reading Data." *Brain and Language*, vol. 120, no. 3, Mar. 2012, pp. 271-81, https://doi.org/10.1016/j.bandl.2011.09.001.

The Man who forever changed Artificial Intelligence. Produced by Sebastian Schuchmann, YouTube, 2019.

https://www.youtube.com/channel/UCf5Rzk7DofjZctBGDzcRVRA/about

McClelland, J. L., et al. "The Appeal of Parallel Distributed Processing." *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations*, e-book ed., pp. 3-44. pdf. https://stanford.edu/~jlmcc/papers/PDP/Chapter1.pdf

Toosi, Amirhosein, et al. "A Brief History of Ai: How to Prevent Another Winter (A Critical Review)." *PET Clinics*, vol. 16, no. 4, pp. 449-69. pdf. Abstract. https://arxiv.org/pdf/2109.01517.pdf

Werbos, Paul J. "Backpropagation Through Time: What It Does and How to Do It." *IEEE*, vol. 78, no. 10, Oct. 1990, pp. 1550-60. pdf. https://axon.cs.byu.edu/Dan/678/papers/Recurrent/Werbos.pdf