Artificial Biology

Artificial vs. Biological Neural Networks

A biological neuron consists of dendrites receiving electrical signals (action potentials) as input, a cell body ('soma') for processing, and an axon for output. An artificial neuron in a feedforward artificial neural network (ANN or 'net'), on the other hand, consists of numerical input, weight, and output values, with processing done in its hidden layers, between the input and output layers. Each hidden node in an ANN transmits numerical signals as the sum of the products of the inputs ($X_1 .. X_n$) and weights ($b_1 .. b_n$) of each node in the previous layer, finally adding the bias (b_0), using multiple linear regression:

$$\hat{Y} = b_1 X_1 + b_2 X_2 + \cdots + b_n X_n + b_0$$

The result is then passed through a non-linear activation function, such as the sigmoid function which outputs in the range [0,1] or the ReLU function which outputs positive values directly or a 0 for negative ones (Gulli, A., 2017). In the output layer, each node in a SoftMax activation (for example) might represent a probability for a classification problem or a probabilistic prediction for a relationship in a regression problem. The predicted output is compared to the ground-truth (correct label or value) using a loss function to determine how far off the network's prediction is. Backpropagation then calculates how much each weight in the net contributed to the error using the calculus chain rule to find each gradient, which is how much the rate of change of the weight (w) affects the Error (E), using its non-linear activation function (a) (Nicholson, C., 2023):

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx} \rightarrow \frac{dE}{dw} = \frac{dE}{da} \cdot \frac{da}{dw}$$

The calculated gradients are then used to update the weights in the direction that reduces error, using an algorithm such as gradient descent (GD). This entire cycle is repeated for many iterations over the training dataset, employing ML techniques along the way such as the following:

- regularization with dropouts to prevent overfitting
- compiling with metrics to observe accuracy, precision, and recall
- trying alternative optimizers (SGD, GD, RMSprop, Adam, etc.)
- optimizing hyperparameters to minimize cost function:
    - number of epochs and hidden neurons, optimizer learning parameters, batch size, etc.
- adjusting number of hidden layers
- augmenting training data to help the model generalize better:
    - rotating, flipping, and adding noise
- GANs (general adversarial networks) with batch normalization to accelerate deep learning and improve accuracy

The most advanced models use deep learning (DL) which utilizes hundreds (or more) layers. As the rate of error is reduced through the evaluation and adjustment cycle, the DL model becomes more accurate in its predictions, involving both simple and complex concepts, as it attempts to approximate an unknown function $f(x) = y$ for any x and y.

In a biological system, the signals from axon terminals (at the tip of each axon) cause neurotransmitters (chemical messengers) to cross the synapse and cause either an excitatory or inhibitory electrical action potential that causes the neuron to either fire or remain dormant (respectively) (Khan Academy, n.d.). Human neural networks are not strictly composed of layers, as each biological neuron can have thousands of complex connections, including recurrent connections and feedback loops. ANNs are simple, high-level representations of their biological counterparts.

Human vs. Machine Learning

When compared to machines, humans do significantly better with one-shot (or few-shot) learning and can learn from small datasets and generalize across domains. ML models, on the other hand, usually require enormous amounts of data and training to understand new concepts. They are easily fooled by added noise and have trouble generalizing beyond specific training data (Fernandez, E., 2019). Machines learn by using mathematical equations to gradually adjust weights for individual nodes in an attempt to learn – this is not how human brains are known to work. Although perhaps some form of complex mathematics is taking place at the cellular or neuronal level.

Both humans and machines are prone to biases in their training and learning. Interestingly, humans and machines can both learn via supervised learning and unsupervised learning and can benefit from adversaries and reward-based reinforcement learning. An example of unsupervised learning for a human is a baby learning to crawl on his or her own, attempting to calculate the input (movement/crawling) that leads to the desired output (mobility). After falling, instead of moving across the floor, the baby realizes something different must be tried! Then, as Mom and Dad try to teach Baby how to control their "legs" and "arms" to move from one place to another, these parents have used a form of supervised learning to teach their child. When Baby begins to move their arms and legs and then finally begin to traverse, Mom and Dad cheer Baby on with varying degrees of enthusiasm, using reinforcement learning to reward the baby with praise and adoration, as well as receiving the reward from traversal itself! As you can see, humans and machines are a lot alike, and perhaps one day, even closer we will be.

References:

Fernandez, Elizabeth. (November 30, 2019). *AI Is Not Similar To Human Intelligence. Thinking So Could Be Dangerous.* Forbes. https://www.forbes.com/sites/fernandezelizabeth/2019/11/30/ai-is-not-similar-to-human-intelligence-thinking-so-could-be-dangerous/?sh=4d78c4ed6c22.

Gulli, Antonio & Pal, Sujit. (2017). Packt Publishing Limited. *Deep Learning with Keras*.

Khan Academy. (n.d.). *How do you know where you are right now?* https://www.khanacademy.org/science/biology/human-biology/neuron-nervous-system/a/overview-of-neuron-structure-and-function.

Nicholson, Chris. (2023). *A.I. Wiki: A Beginner's Guide to Neural Networks and Deep Learning*. Pathmind. https://wiki.pathmind.com/neural-network.