Deep Learning Note

Muhamad Ridzuan

March 5, 2024

1 Chapter 1

1.1 Donald Hebb

Organization of behaviour - 1949 learning mechanism:

- When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as on of the cells firing B, is increased.
- As A repeatedly excites B, its ability to excite B improves.
- Neuron that fire together wire together.

1.2 Hebbian Learning

- If neuron x repeatedly triggers neuron y, the synaptic knob connecting x to y get larger.
- In mathematical model:
$$w_{xy} = w_{xy} + \eta xy$$
- Weight of the connection from input neuron x to output neuron y.
- This simple formula is actually the basis of many learning algorithms in machine learning.

This idea however is fundamentally unstable:

- Stringer connections will enforce themselves.
- No notion of "competition".
- No *reduction* in weights.
- Learning is unbounded.p

People came up with all kinds of modifications for it to try to make it more stable:

- Allowing for weight normalization.
- Forgetting

This lead to the Generalized Hebbian learning, aka Sanger's rule where the contribution of input is *incrementally distributed* over multiple outputs.

$$w_{ij} = w_{ij} + \eta y_j \left(x_i - \sum_{k=1}^{j} w_{ik} y_k \right)$$

1.3 A better model

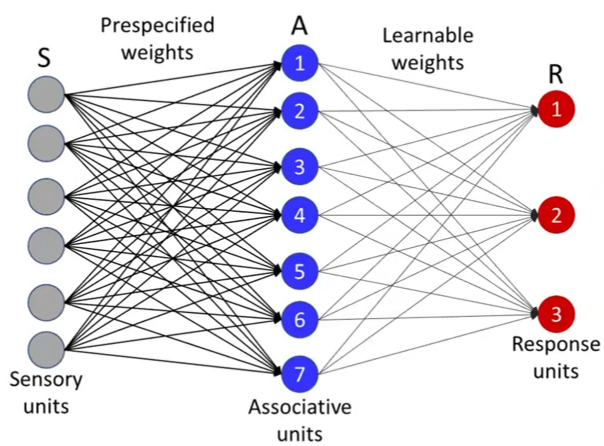
Frank Rosenblatt

- Psychologist, Logician
- Inventor of the solution to everything, aka Perceptron

Original perceptron model Consider the eye structure:

- Groups of sensors on retina combine into cells in association in the **projection area**.
- Groups of **projection area** combine into Association cells in **association area**.
- Signals from **association area** cells combine into response cell **R**.
- All connectioons may be excitatory or inhibitory.

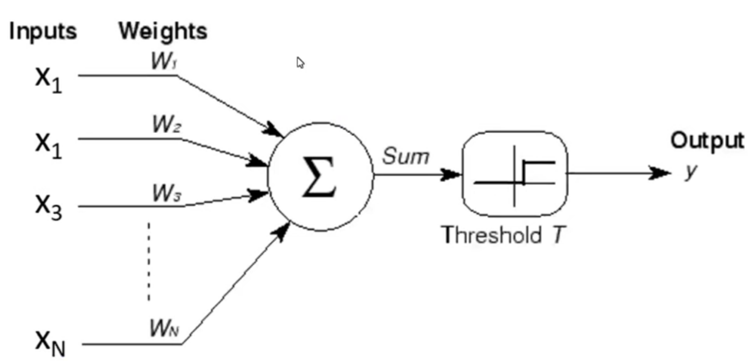
Rosenblatt's perception model can then be further simplified:



Simplified perceptron model

- Association units commbine sensory input with fixed weights.
- Response units combine associative units with learnable weights.

Each of the units can be perceived as shown below:



- Number of inputs combine linearly.

- Threshold logic is applied at the output of the unit: It will only fire if the weighted sum of the input exceeds threshold.

$$Y = \begin{cases} 1, & \text{if } \sum_i w_i x_i - T > \ 0 \\ 0, & \text{else} \end{cases}$$

1.4 The Universal Model

Originally assumed could represent any Boolean circuit and perform any logic. However this was not true and this cause the research in neural networks died down because of the overhype. However, Rosenblatt was right, he not only gave us basic model, he also gives us the learning algorithm.

$$w = w + \eta(d(x) - y(x))x$$

Sequential learning:

- $d(x)$ is the desired output in response to input x.

- $y(x)$ is the actual output in response to x.

- Weight parameter is changed when the actual response of the neuron does not match the desired response of the neuron.

- If the actual response of the neuron exceeds the desired response of the neeuron, then the weight will decrease.

- If the actual response is lesser than the desired response, the weight will increase.