Parceptron

Akira Imada

Most recently renewed on

March 22, 2010

1 Single Parceptron

Assume we have 10 inputs $x_1, x_2, \cdots, x_{10}$ and one output y. Also assume each input is binary 1 or -1. All the input is connected to output with synapse whose weight is w_i ($i = 1, 2, \cdots 10$). If $w_1x_1 + w_2x_2 + \cdots + w_{10}x_{10}$ is larger than a threshold θ then $y = 1$, otherwise $y = -1$. That is,

$$y = sgn(\sum_{i=1}^{10} w_i x_i - \theta). \quad (1)$$

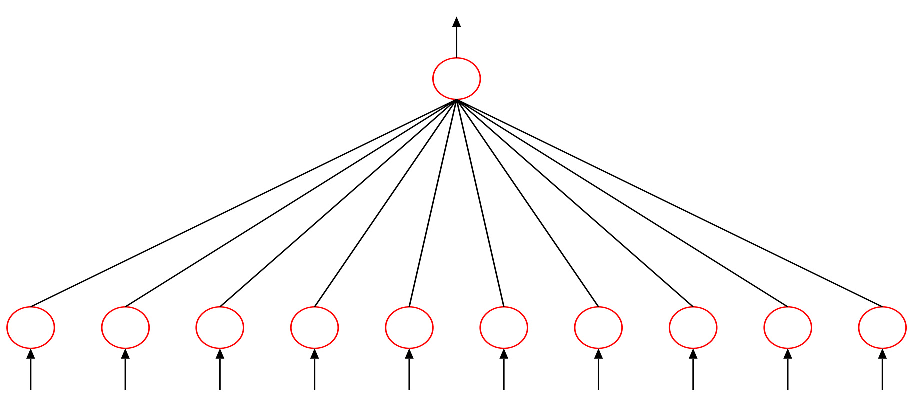

Figure 1: Schematic diagram of Single Parceptron.

Task 1 (A Framework) *Assigning random weight value from -1 to 1, create a program which accepts M binary inputs from keyboard, calculate y, and display 10 input and output with their color being green if the value is 1 or red if the value is -1. Also show weighted sum of the inputs $\sum_{i=1}^{10} w_i x_i$ on the screen like the following Figure.*

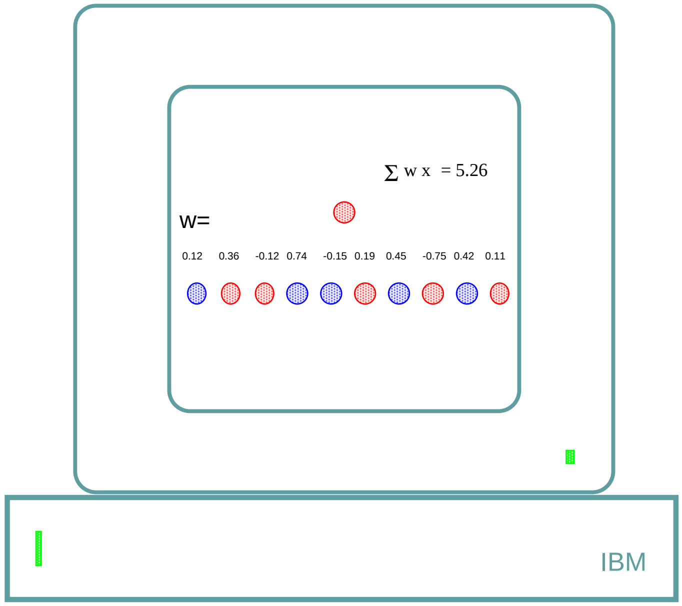

Figure 2: An example of display of the task.

Examples of how can it work.

One of the most typical example is to realise a Boolean function. For example, output is 1 if and only if all the input is 1, otherwise -1.

Table 1: An example of Boolean Function assuming 4 inputs case. Note that we call this AND logic when the number of input is 2.

x_1	x_2	x_3	x_4	x_5	x_6	x_7	y
-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	1	-1
-1	-1	-1	-1	-1	-1	1	-1
...	...	...	...	...	...	...	...
1	1	1	1	1	1	1	-1
1	1	1	1	1	1	1	1

Task 2 (Extended AND) *Discover a set of weights and threshold so that $y = 1$ if and only if input are all 1, otherwise $y = -1$.*

Task 3 (Extended OR) *Discover a set of weights and threshold so that $y = 1$ if at least one input is 1 and $y = -1$ if all input is -1.*

Task 4 (EVEN PARITY) *What about the case where $y = 1$ if the number of 1 is odd, otherwise $y = -1$?*

2 Perceptron Learning

Assume now we have m input and n output w_{ij} is a weight value from input j to output neuron i. X^p is p-th input vector for training, that is,

$$X^p = (x_1^p, x_2^p, x_3^p, ...x_n^p),$$

$\hat{Y}^p$ is target output vector when p-th training input X^p is given, that is,

$$\hat{Y}^p = (\hat{y}_1^p, \hat{y}_2^p, \hat{y}_3^p, \cdots, \hat{y}_m^p)$$

Y^p is actual output vector, that is,

$$Y^p = (y_1^p, y_1^p, y_1^p, \cdots y_m^p)$$

Then we can describe a learning of weights as

$$w_{ij}(t+1) = w_{ij}(t) + \eta(\hat{Y}^p - Y^p)(X^p)^t.$$

Note that $(\hat{Y}^p - Y^p)(X^p)^t$ is an inner product of two vectors, that is,

$$(\hat{y}_1^p - y_1^p)x_1^p + (\hat{y}_2^p - y_2^p)x_2^p + (\hat{y}_3^p - y_3^p)x_3^p + \cdots + (\hat{y}_m^p - y_m^p)x_m^p.$$

η is a learning coefficient set to a small real number ranging $(0 < \eta < 1$. The learning is repeated until the change in weight value becomes neglectable. In other words, we have to preset p enough a large number so that $w_{ij}(t+1) - w_{ij}(t+1)$ becomes almost 0, after repeating above $p$times. This is also called *Widrow-Hoff Learning Algorithm*

3 Multi-layer Parceptron

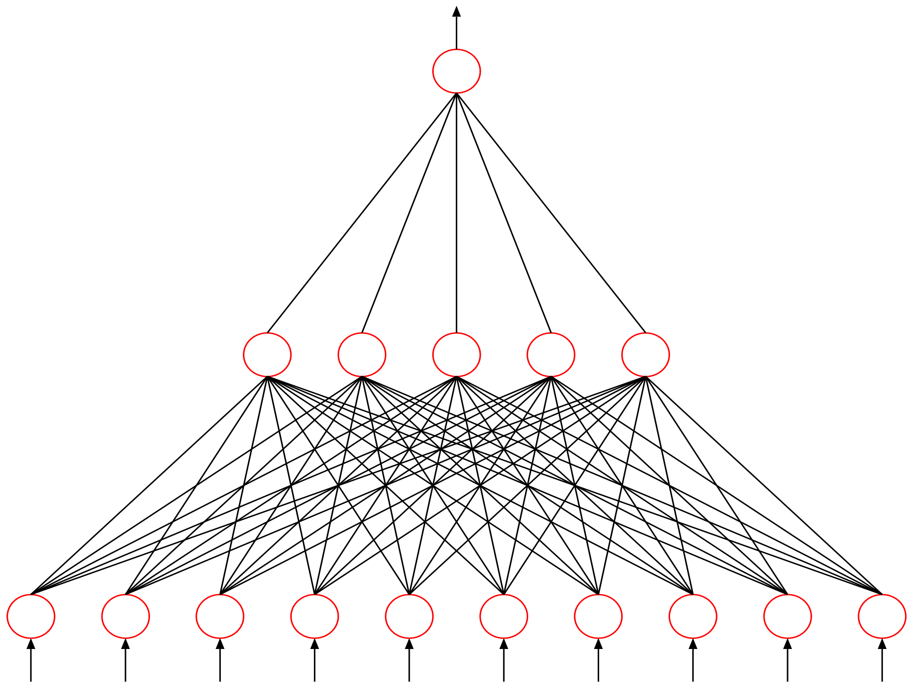

Figure 3: Schematic diagram of Multi-layer Parceptron.