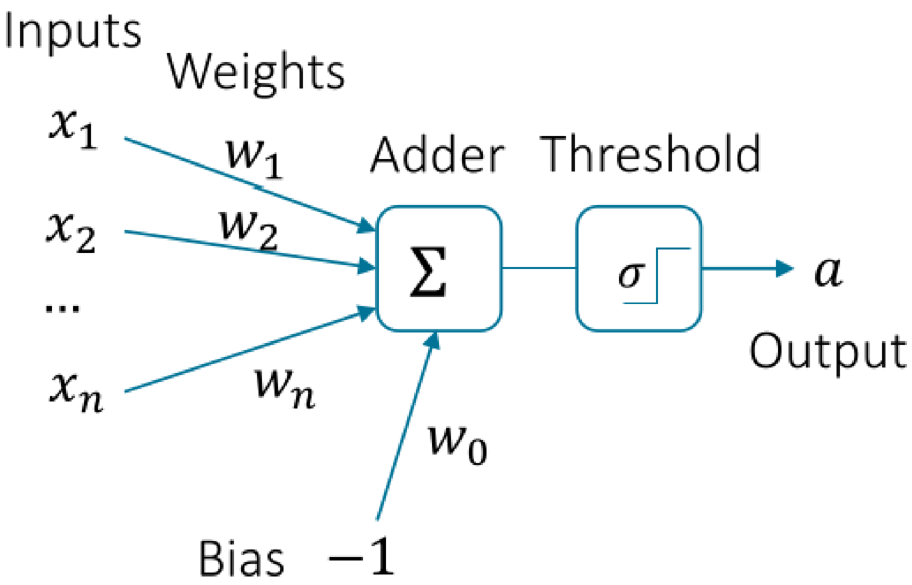
DTE-2502: Neural Networks
Module01: Perceptron
Basic definitions
Kalyan Ram Ayyalasomayajula, email: kay001@uit.no

1. Neuron model

Given n numerical features x_j, $j = \{1, \cdots, n\}$ $a(x, w) = \sigma(\langle w, x\rangle) = \sigma(\sum_{j=1}^{n} w_j x_j - w_0)$, where $w_0, w_1, \cdots, w_n \in \mathbb{R}$ are feature weights, $\sigma(z)$ is an activation function, for example, $sign(z) = \begin{cases} +1, z \leq 0 \\ -1, z < 0 \end{cases}$, $sigmoid(z) = \frac{1}{1-e^{-z}}$, etc.



Linear neuron model
(McCulloch and Pitts, 1943)

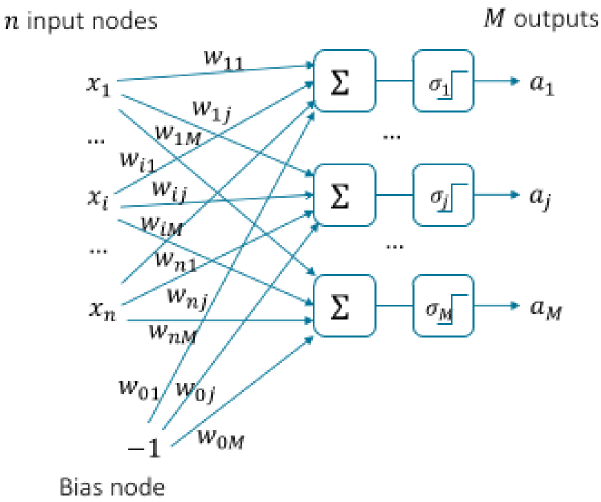
2. Perceptron model

Definition 1.
The perceptron is a collection of neurons together with a set of inputs and weights to associate the inputs to the neurons.

- Weights $w_{ij}, i = \{1, \cdots, n\}, j = \{1, \cdots, M\}$
- The rule for updating a weight w_{ij}

$$w_{ij} = w_{ij} - \mu(a_j - y_j) \cdot x_i \tag{1}$$

, where y_j is the target value, $0.1 \leq \mu \leq 0.4$ is the learning rate.



Definition 2: The perceptron learning algorithm.

Initialization

Set all the weights w_{ij} to small positive and negative random values

Training phase

for T iterations or until all the outputs are correct

for each input vector

compute the activation of each neuron using *activation function* σ

$$a_j = \sigma\left(\sum_{i=0}^{M} w_{ij}x_i\right) = \begin{cases} 1 & \text{if } \sum_{i=0}^{M} w_{ij}x_i > 0 \\ 0 & \text{if } \sum_{i=0}^{M} w_{ij}x_i \leq 0 \end{cases}$$

update each of the weights individually using

$$w_{ij} \coloneqq w_{ij} - \mu\left(a_j - y_j\right) \cdot x_i$$

Recall phase

compute the activation of each neuron as $a_j = \sigma\left(\sum_{i=0}^{M} w_{ij}x_i\right) = \begin{cases} 1 & \text{if } \sum_{i=0}^{M} w_{ij}x_i > 0 \\ 0 & \text{if } \sum_{i=0}^{M} w_{ij}x_i \leq 0 \end{cases}$