Topic 7.4 Neural Network Learning

A) An abstract example of a simple artificial Neural Network

$x_1 \longrightarrow$ **1**

$x_2 \longrightarrow$ **2**

...

$x_i \longrightarrow$ **i** w_{ij} **j** w_{jk} **k** $\longrightarrow O_k$

Input layer   Hidden layer   Output layer
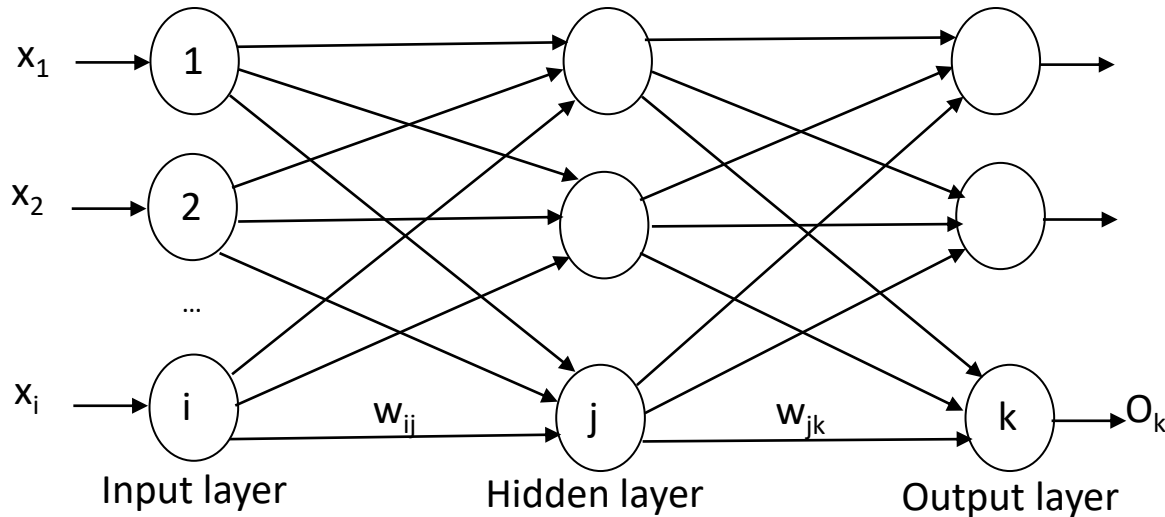
Two-layer fully connected feed-forward Neural Network

1. (x_1, x_2, ... , x_i) – numerically scaled and normalized attribute values of a sample

2. The input layer only copies

3. May be more than one hidden layers

4. Connections are weighted

x_1 → 1

x_2 → 2

…

x_i → i w_{ij} j w_{jk} k → O_k

Input layer   Hidden layer   Output layer
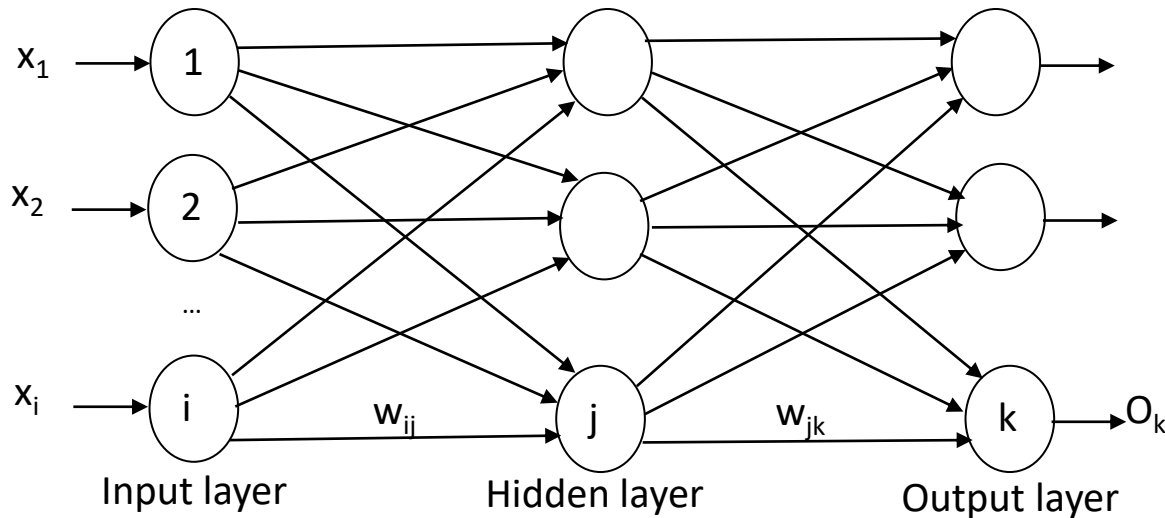
5. Weighted output of one layer is passed on to the next.

6. Simply, a Neural Network is a set of connected input/ output units, where each connection has a weight associated with it.

7. A unit in hidden layer or output layer is called a neurode (neuron like) or output unit.

8. Training Samples are fed and network parameters like weights are adjusted based on feedback (the last layer output).

9. Initial topology (layers, units, connectivity, …) and set of parameters like weights affect the performance, and can be reset if can't be trained up to the mark.

10. Various types of networks are possible: Acyclic or feed forward; Cyclic or recurrent; Single layer (hidden) / Multilayer.

B. A closer look into a neurode, j:
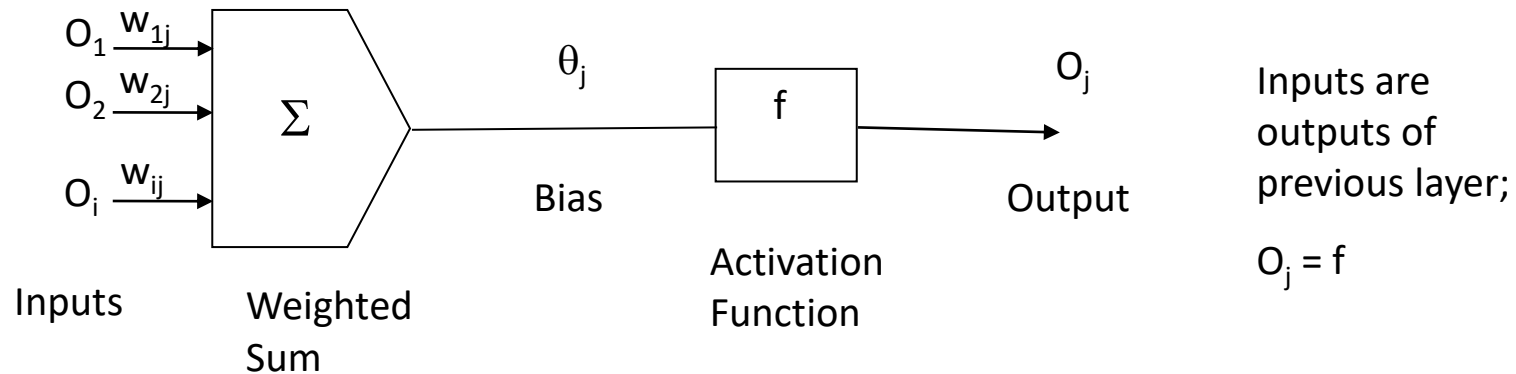
- θ_j : to vary the influence; threshold to determine whether to fire or not

- Net input, $I_j = x = \Sigma_i\, O_i * w_{ij} + \theta_j$ [O_i- output of a previous layer unit]

- $O_j = 1 / (1+e^{-x})$
  - ✓ Squashing function (maps large domain to [0, 1])
  - ✓ Nonlinear
  - ✓ Sigmoid function (S-shaped) / Logistic curve ('population' growth rate); Widely used in statistics

C. Features of Neural Network Learning

1. It is a kind of <mark>statistical learning</mark> methods (finding predictive function from field data).

2. Basic idea: <u>Models</u> information processing aspects of <u>neurons</u>.

3. Learning Phase: Adjusting parameters like weights and biases, so as to be able to predict <u>correct</u> class of an input.

4. <mark>Disadvantages:</mark> Long training time; Topology, weights, etc., are determined empirically; Poor interpretability.

5. <mark>Advantages:</mark> High tolerance to noisy data; High ability to classify patterns.

6. <mark>Backpropagation</mark> is the most popular algorithm for NNL.