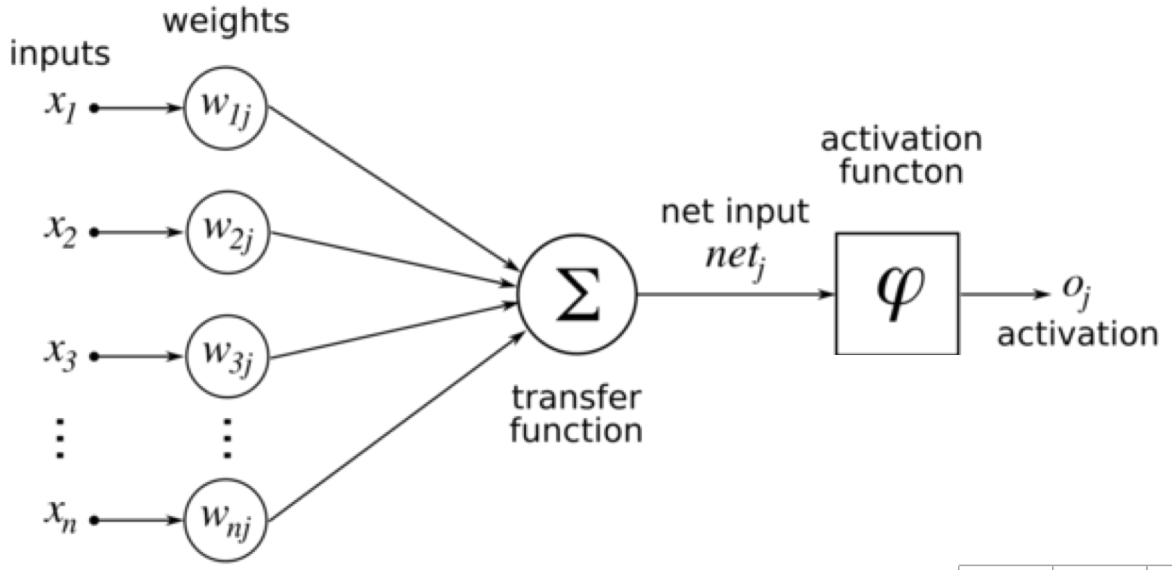
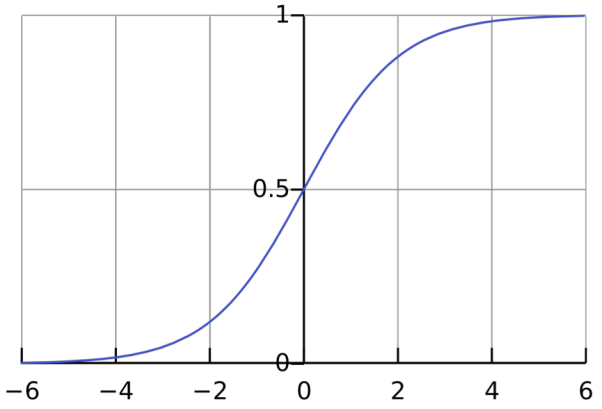
Back-Prop

An artificial Neuron



Logistic activation function: $\varphi(x) = \dfrac{1}{1+e^{-x}}$

Derivative: $\dfrac{d}{dx}\varphi(x) = \dfrac{e^{-x}}{\left(1+e^{-x}\right)^2} = \varphi(x)\left(1-\varphi(x)\right)$

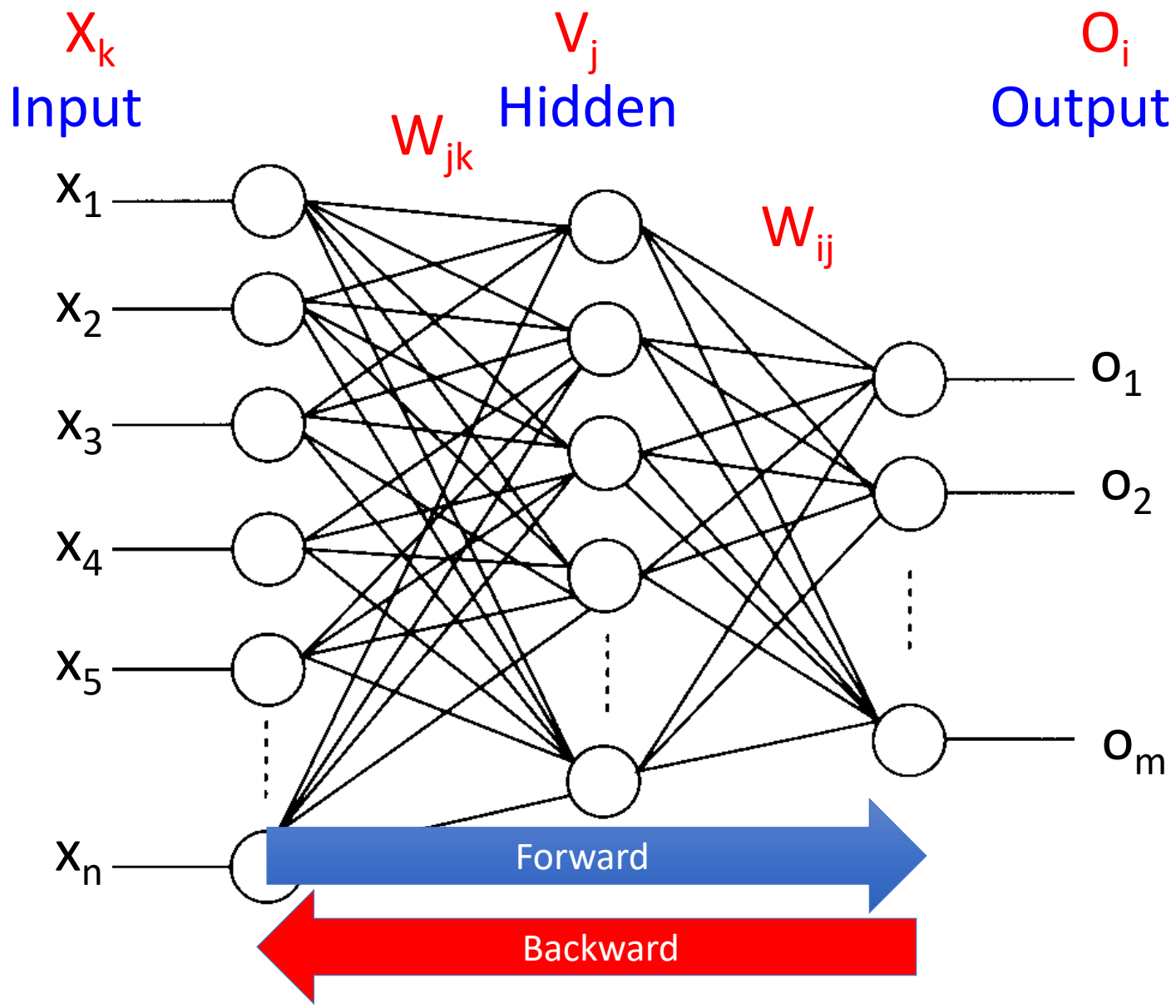


Learning update of a single neuron

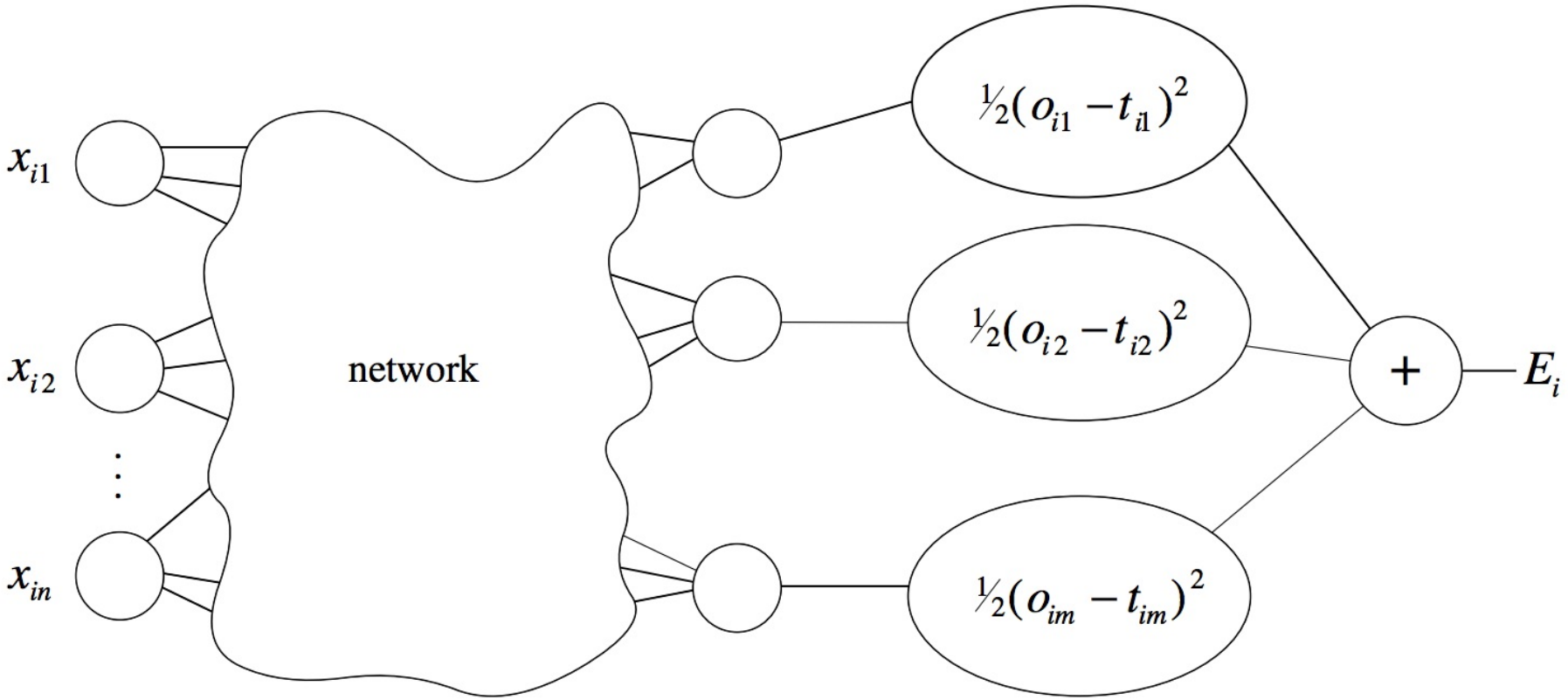
A source of data: $(\vec{x}_1, y_1), (\vec{x}_2, y_2), \ldots$ where $\vec{x}_t \in R^n, y_t \in [0,1]$

A feed-forward Artificial Neural Network (ANN)

Based on “Introduction to the theory of Neural computation” Hertz, Krogh and Palmer

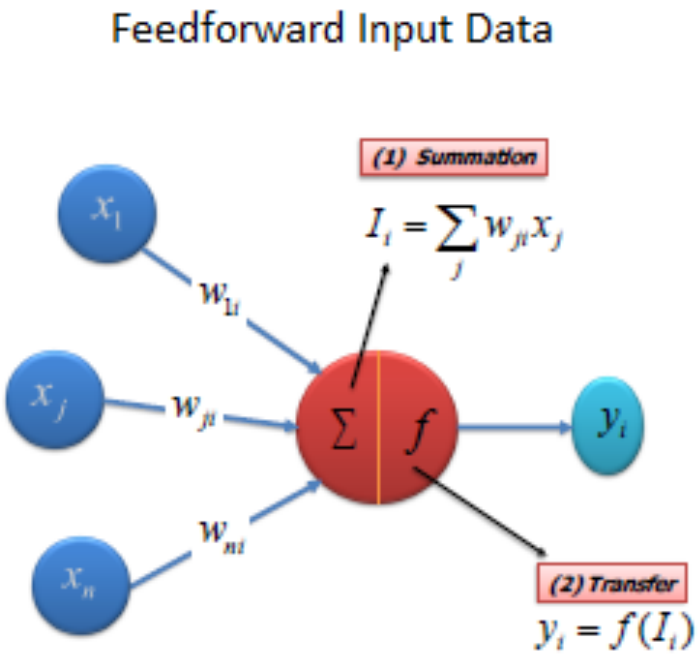


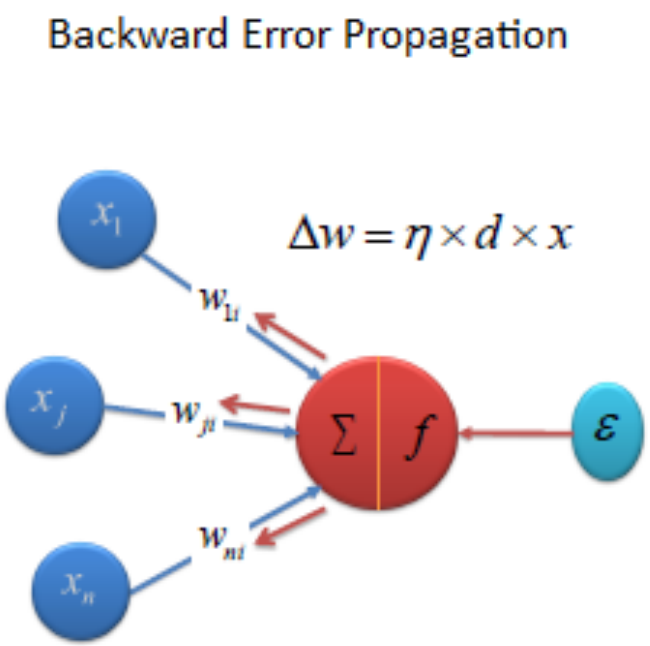
ANN loss function



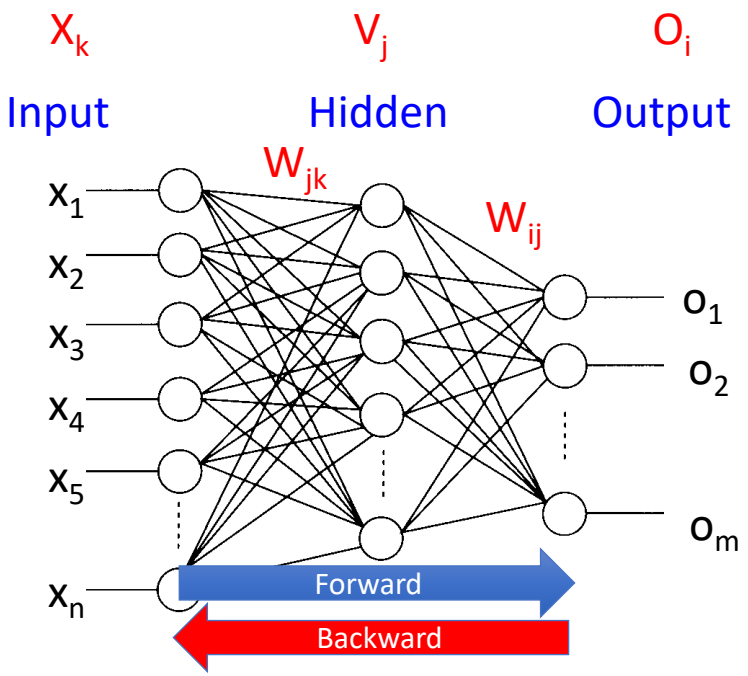
Computing the gradient

- Forward pass: propagate activations.
- Backward pass: propagate errors.





Forward-Backward
Update rules



Summary

- Most NN are feed-forward
  - There are NN with Feedback, we will not talk about them here.
- A Layered FF-NN has connection going from one layer to the next.
- **The forward pass** computes the activation level of all neurons/
- Given the output we compute the **errors** for output nodes.
- **The backwards pass** computes the effective error for each internal neuron.
- In the backward pass the weights are updated.