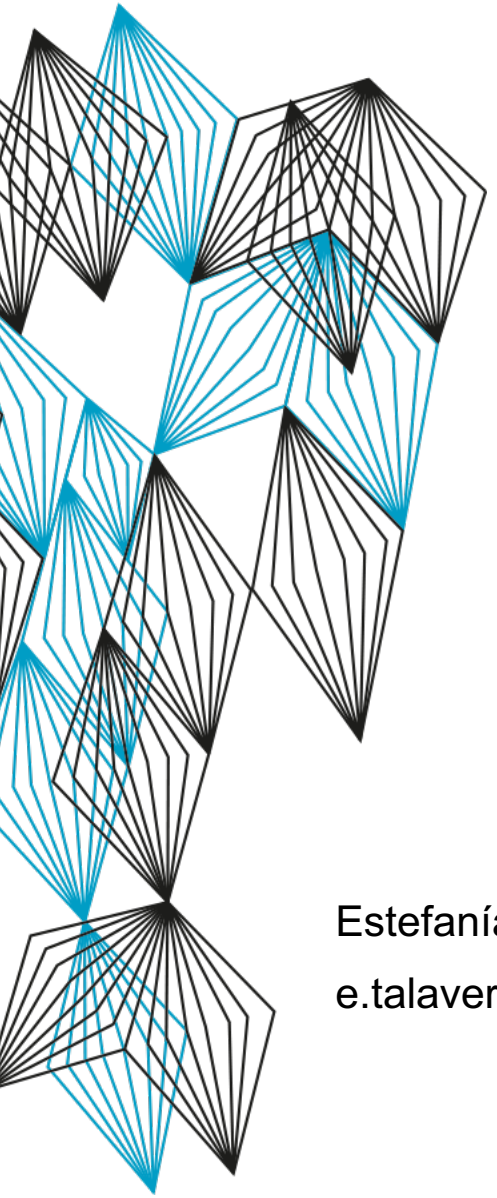


ARTIFICIAL INTELLIGENCE & CYBER SECURITY

NETWORKS NETWORKS

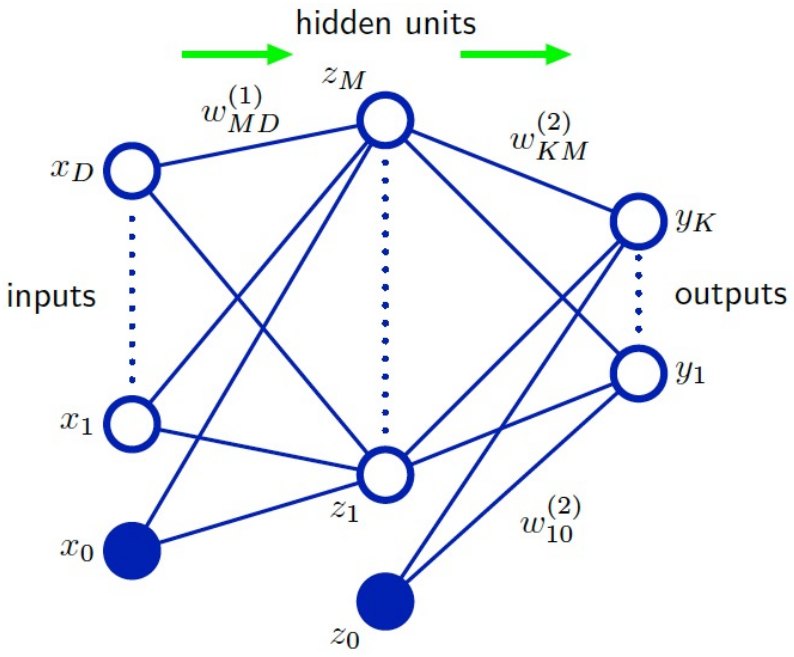
Multilayer perceptron

Estefanía Talavera Martínez

e.talaveramartinez@utwente.nl

MULTI-LAYER PERCEPTRONS

With a clever application of the chain rule of derivations we can combine multiple layers and still train the network.



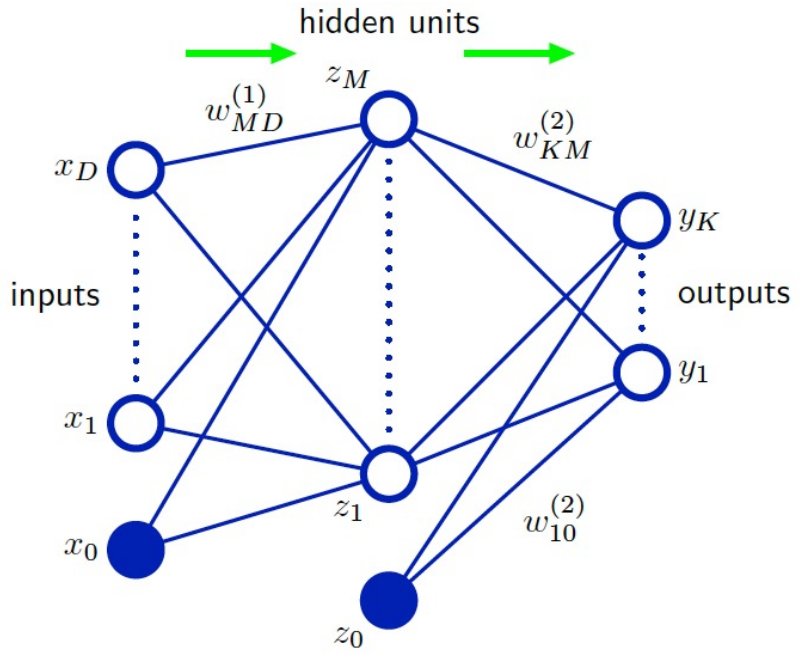
The architecture is constrained

- In order to be trainable, a feed-forward architecture is required
- Can be sparse
- Can have skip-layer connections

This is clearly much more constrained than biological neural networks

MULTI-LAYER PERCEPTRONS

UNIVERSAL FUNCTION APPROXIMATORS



The combined, weighted non-linearities make very complex functions possible

A two-layer network with “linear" output activation function can approximate any continuous function within a compact domain with arbitrary precision

- If the hidden layer has sufficient units
- Holds for many activation functions of the hidden units (but not polynomials)

But! Deep networks are often more compact and effective than wide networks!

Combining two layers results in function of the form

$$y_k(\mathbf{x}, \mathbf{w}) = h_2\left(\sum_{j=0}^{M} w_{kj}^{(2)} h_1\left(\sum_{i=0}^{D} w_{ji}^{(1)} x_i\right)\right)$$