Multi-layer Perceptron Architecture

Input Layer

x_1

x_2

x_3

x_4

n features

Hidden 1

h_1^1

h_2^1

h_3^1

h_4^1

h_5^1

m_1 units

Hidden 2

h_1^2

h_2^2

h_3^2

h_4^2

m_2 units

Hidden 3

h_1^3

h_2^3

h_3^3

m_3 units

Output Layer

$\hat{y}_1$

$\hat{y}_2$

k outputs

Layers
L = 5

Hidden Layers
3 layers

Parameters
Weights + Biases

Connectivity
Fully Connected

$$h^{(l)} = \sigma(W^{(l)}h^{(l-1)} + b^{(l)}) \rightarrow \text{Layer-wise Computation}$$