

DEEP LEARNING

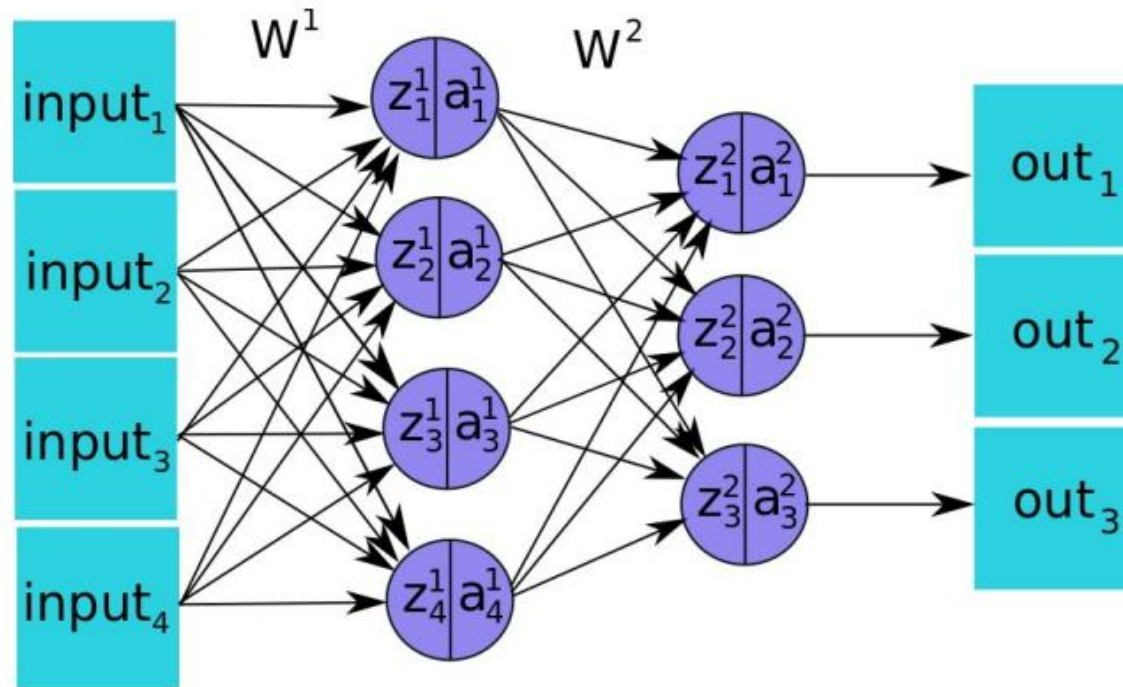
Сеть прямого распространения

ГУУ, 3-й курс 2023, 2-й семестр

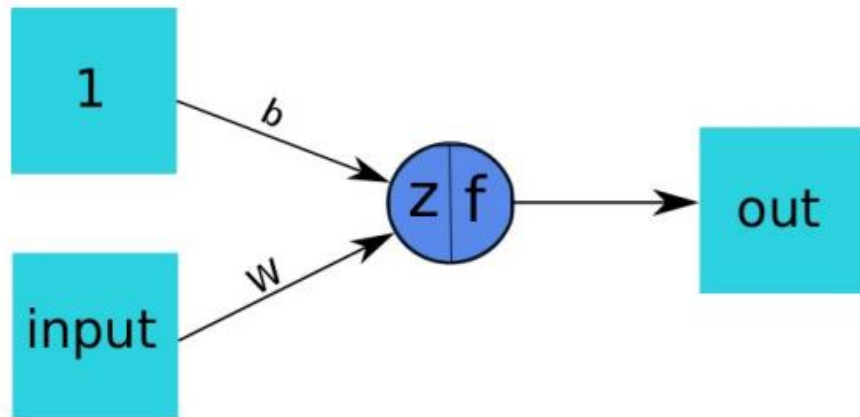
ПЛАН

- Сеть прямого распространения
- Функция активации
- Алгоритм обратного распространения
- Автоматический градиент

НС прямого распространения

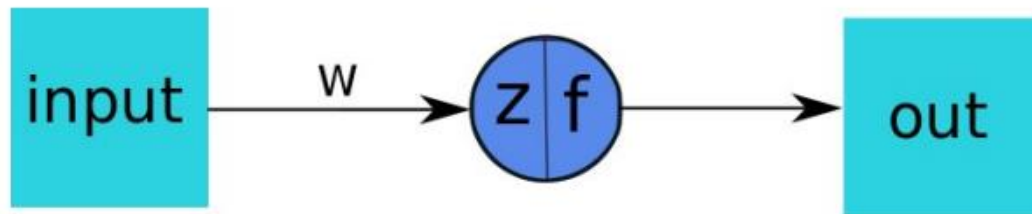


Один вход, один нейрон



$$out = f(w \cdot input + 1 \cdot b)$$

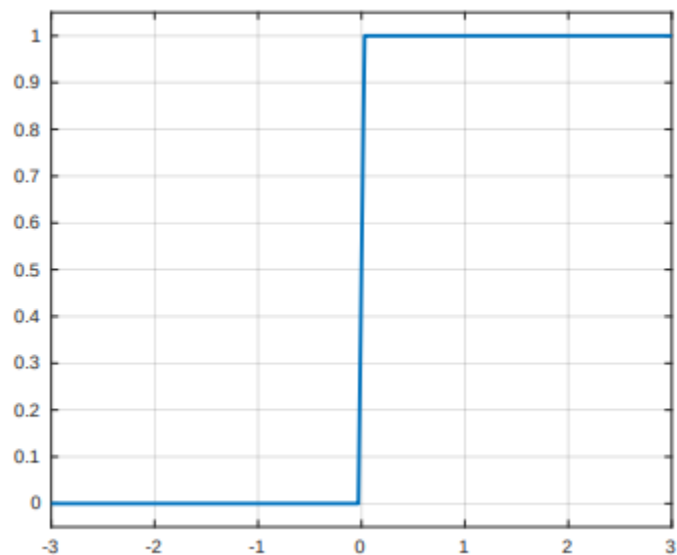
$$z = w \cdot input + 1 \cdot b$$



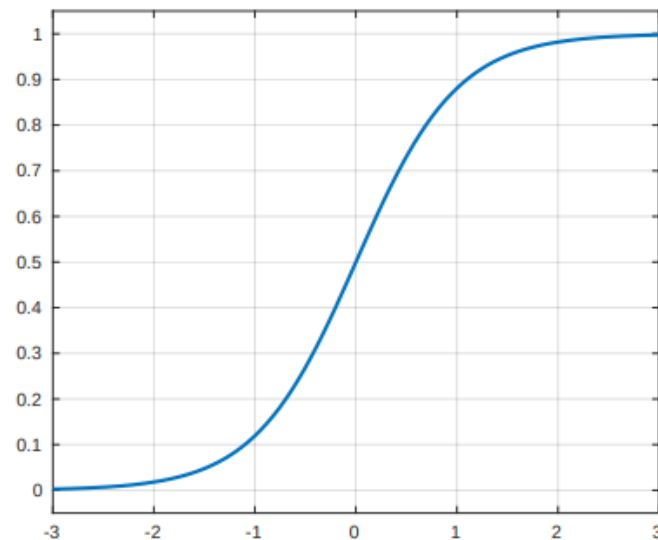
$$out = f(z)$$

Функция активации

$$f(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ -1, & \text{if } x < 0 \end{cases}$$

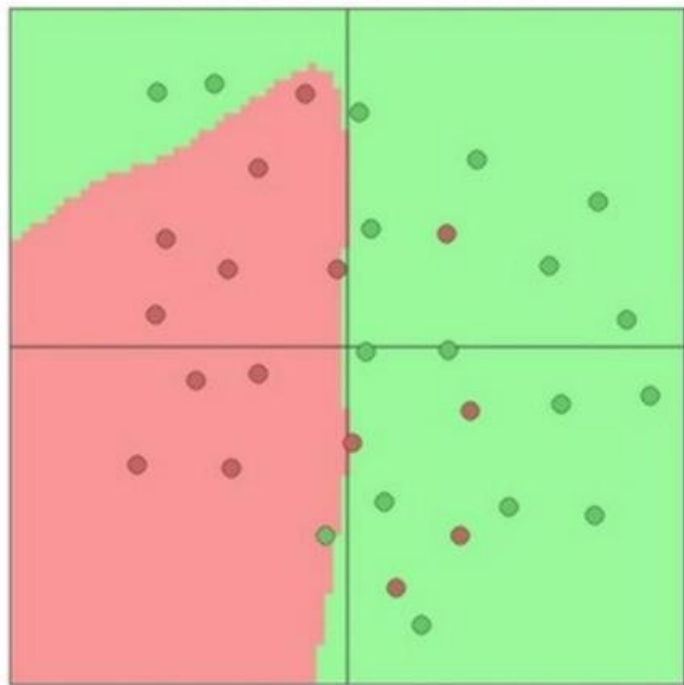


$$f(x) = \frac{1}{1 + e^{-x}}$$

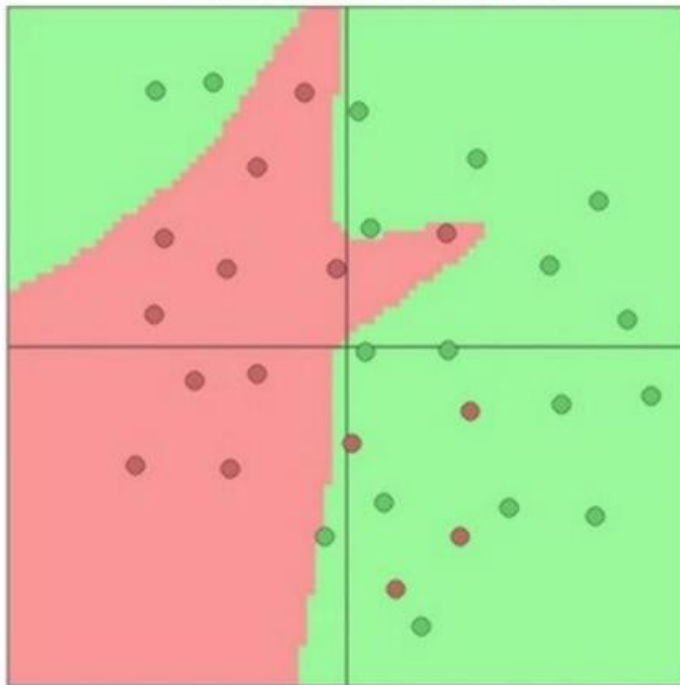


Функция активации

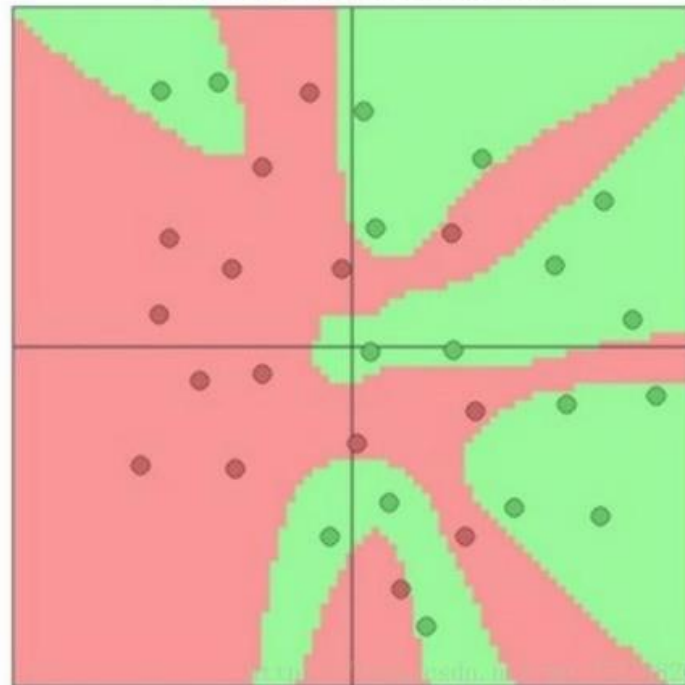
3 hidden neurons



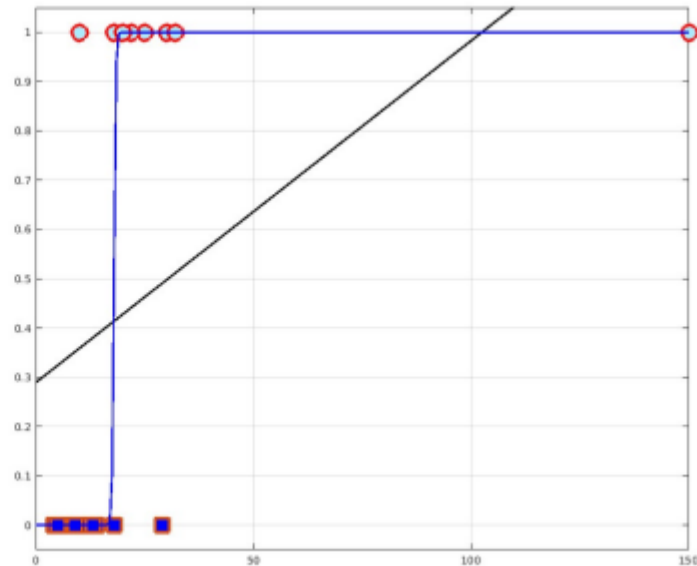
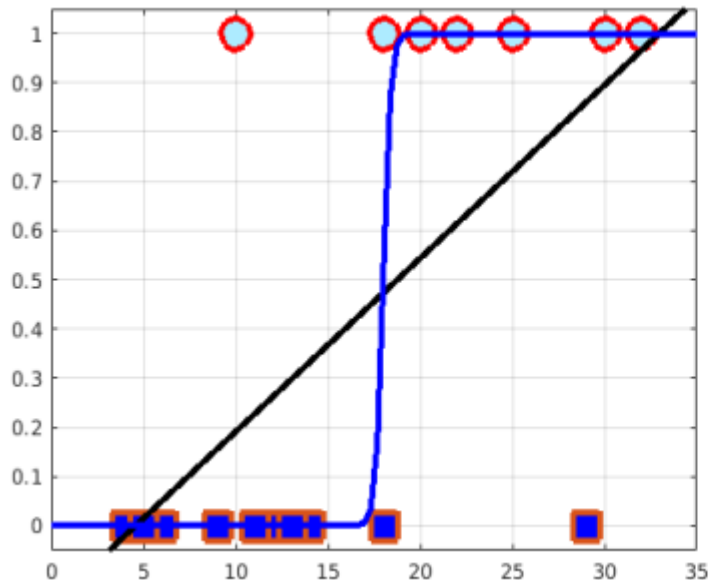
6 hidden neurons



20 hidden neurons

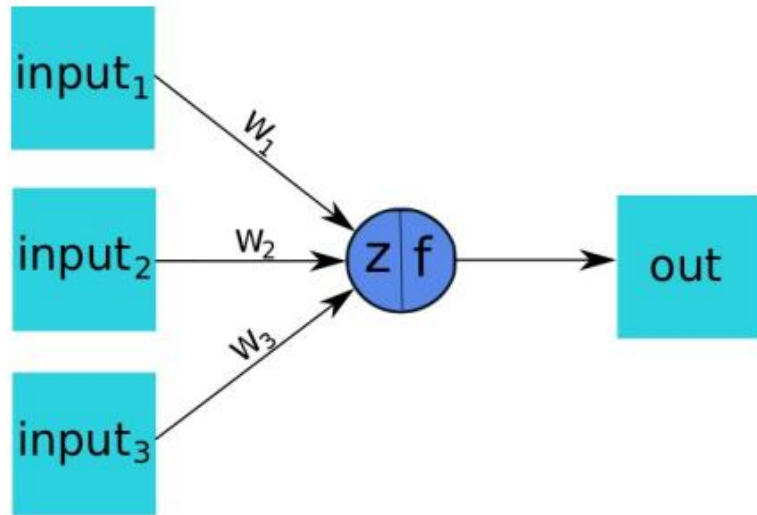


Функция активации: СИГМОИД



$$f(z) = \sigma(z) = \frac{1}{1 + e^{-z}}$$

НС: 3 входа, 1 скрытый нейрон

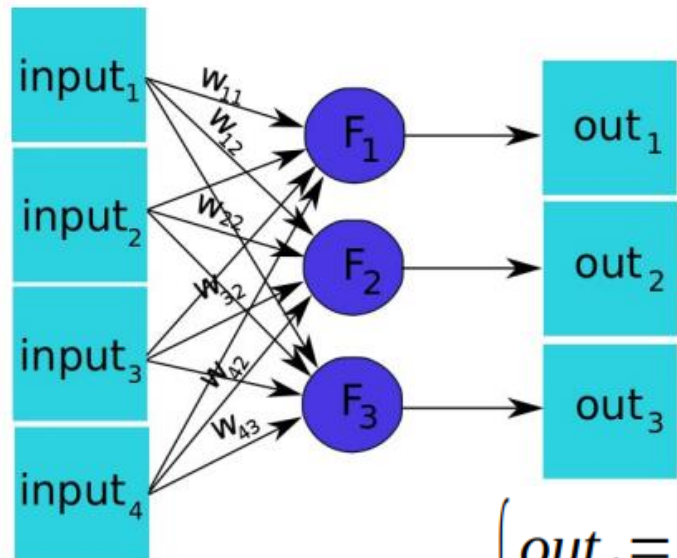


$$out = f(z)$$

$$z = w_1 \cdot input_1 + w_2 \cdot input_2 + w_3 \cdot input_3 + b$$

$$f(z) = \frac{1}{1 + e^{-z}} = \frac{1}{1 + e^{-(w_1 \cdot input_1 + w_2 \cdot input_2 + w_3 \cdot input_3 + b)}}$$

НС: 4 входа, 3 выхода



$$INPUT = \begin{pmatrix} input_1 \\ input_2 \\ input_3 \\ input_4 \end{pmatrix}$$

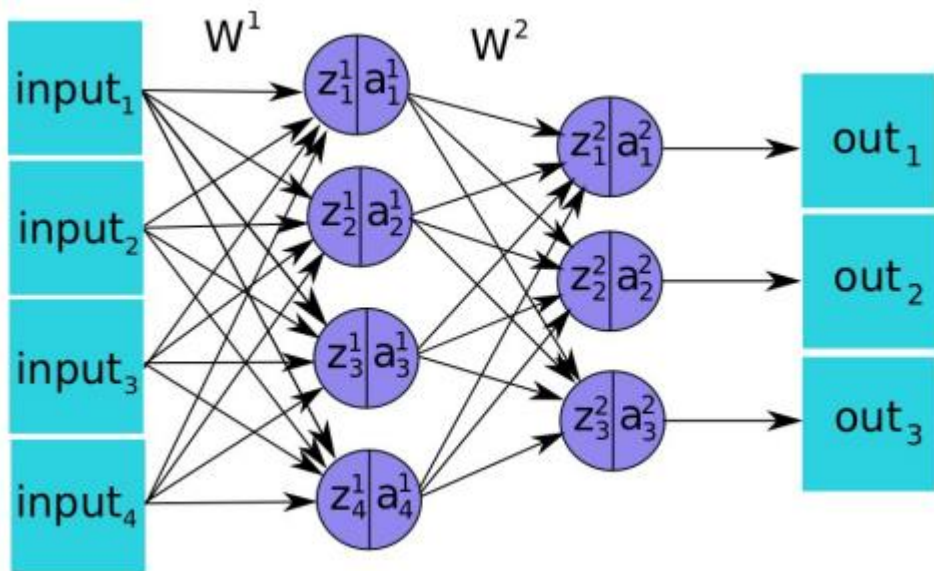
$$B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

$$W = \begin{pmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \\ w_{41} & w_{42} & w_{43} \end{pmatrix}$$

$$OUT = F(W^T \cdot INPUT + B)$$

$$\begin{cases} out_1 = F(w_{11} \cdot input_1 + w_{21} \cdot input_2 + w_{31} \cdot input_3 + w_{41} \cdot input_4 + b_1) \\ out_2 = F(w_{12} \cdot input_1 + w_{22} \cdot input_2 + w_{32} \cdot input_3 + w_{42} \cdot input_4 + b_2) \\ out_3 = F(w_{13} \cdot input_1 + w_{23} \cdot input_2 + w_{33} \cdot input_3 + w_{43} \cdot input_4 + b_3) \end{cases}$$

Многослойная сеть



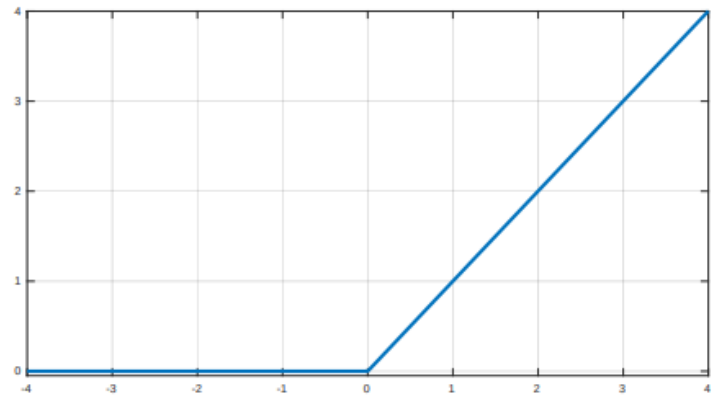
$$OUT^1 = A(W^{1T} \cdot INPUT + B^1)$$

$$OUT = A(W^{2T} \cdot INPUT^2 + B^2) = A(W^{2T} \cdot OUT^1 + B^2)$$

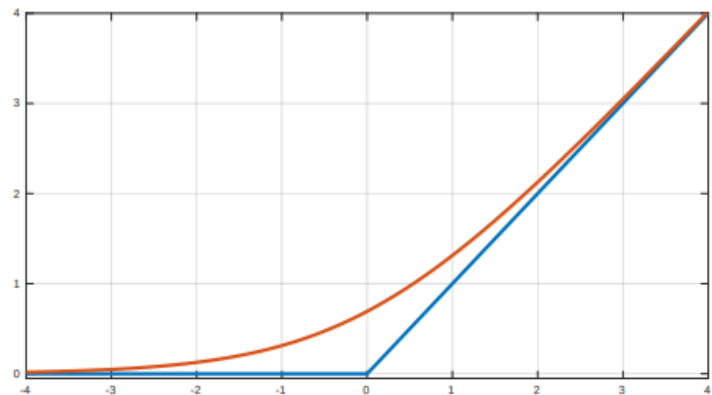
$$OUT = A(W^{2T} \cdot A(W^{1T} \cdot INPUT + B^1) + B^2)$$

Ректификация ReLU

$$y = \max(x, 0)$$



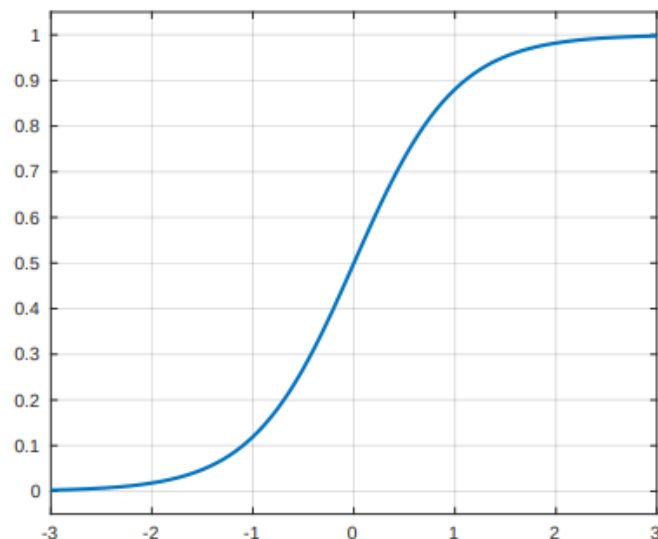
$$y = \ln(1 + e^x)$$



Ректификация ReLU

$$f(x) = \frac{1}{1 + e^{-x}}$$

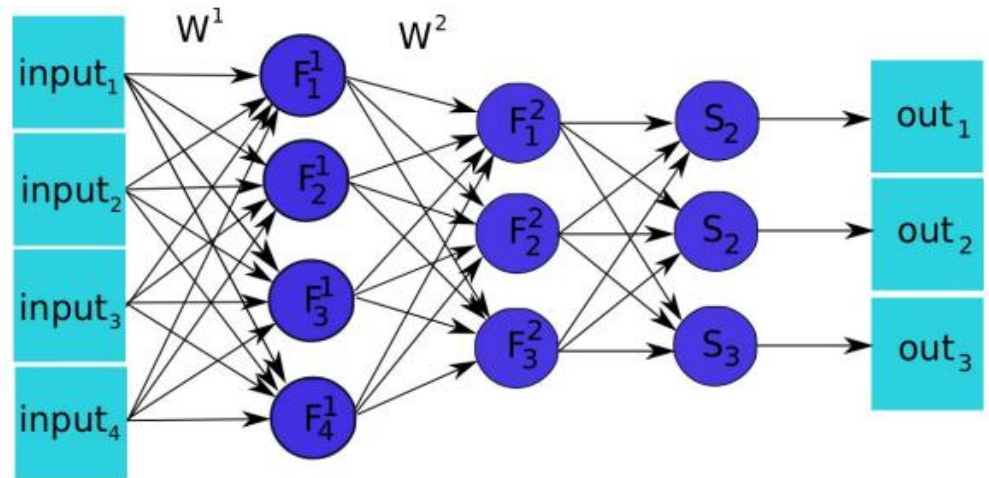
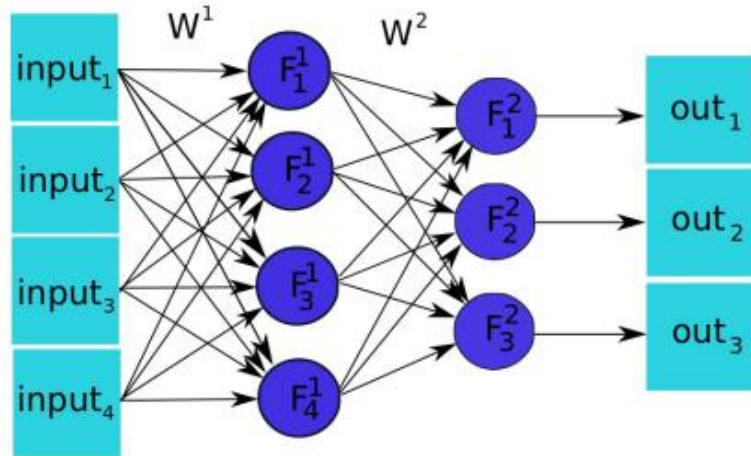
$$f(5) \approx f(10)$$



$$f(x) = \sigma\left(x + \frac{1}{2}\right) + \sigma\left(x - \frac{1}{2}\right) + \sigma\left(x - \frac{3}{2}\right) + \sigma\left(x - \frac{5}{2}\right) + \dots$$

$$f(x) = \sum_{i=0}^{\infty} \sigma\left(x + \frac{1}{2} - i\right) \approx \int_{1/2}^{\infty} \sigma\left(x + \frac{1}{2} - y\right) dy = \left[-\log\left(1 + e^{x + \frac{1}{2} - y}\right)\right]_{y=1/2}^{y=\infty} = \log(1 + e^x)$$

Софтмакс



$$P(C_j|data) = \frac{P(data|C_j)P(C_j)}{\sum_{k=1}^K P(data|C_k)P(C_k)}$$

$$z_k = \ln P(data|C_k)P(C_k)$$

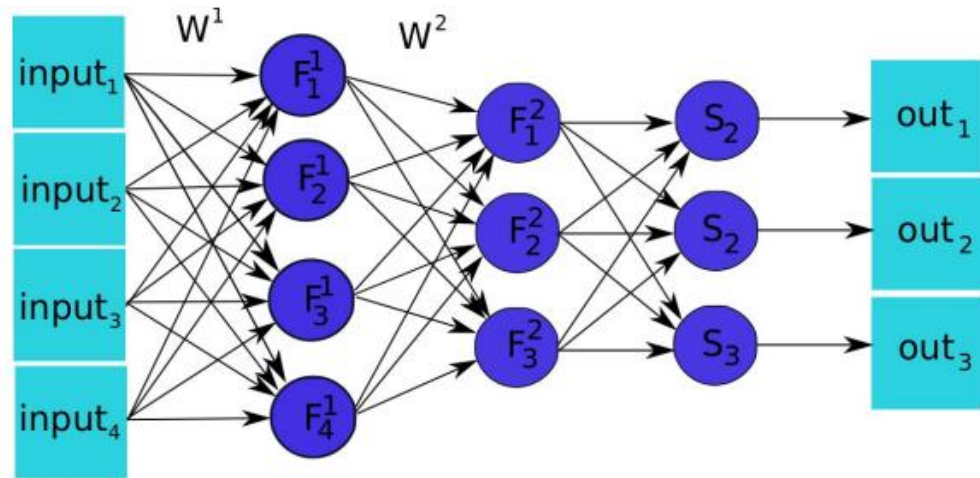
$$\sigma(z_j) = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}$$

Принцип максимального правдоподобия

$$p(C = out | input, w, b)$$

$$p(data) = \prod_i p(C = out_i | input_i)$$

$$-\ln p(data) = -\sum_i \ln(C = out_i | input_i)$$



Функция ошибки

Регрессия

$$E = \sqrt{Y^2 - T^2}$$

Классификация

$$H(p, q) = H(p) + D_{KL}(p, q)$$

$$H(p, q) = - \sum_x p(x) \ln(q(x))$$