Московский государственный университет имени М.В. Ломоносова
физический факультет
NeuroFuzzy Lab.

Искусственные нейронные сети (ИНС)

Зубюк Андрей Владимирович
zubjuk@physics.msu.ru

http://NeuroFuzzy.Phys.MSU.ru

МФК «Искусственный интеллект в естествознании»

Искусственная нейронная сеть

Определение

Искусственная нейронная сеть — функция $F^{(L)}_{\text{ANN}} : \mathcal{R}_n \to \mathcal{R}_m$, представимая рекуррентной формулой

$$F^{(0)}_{\text{ANN}}(x) = W^{(0)}x + b^{(0)}$$

$$F^{(i)}_{\text{ANN}}(x) = W^{(i)} f_i\left(F^{(i-1)}_{\text{ANN}}(x)\right) + b^{(i)}, \quad i = 1, \ldots, L$$

- $\mathcal{R}_{N_i}$ — евклидово пространство столбцов высоты N_i, $i = 0, \ldots, L+1$,
- $W^{(i)}$ — линейный оператор $\mathcal{R}_{N_i} \to \mathcal{R}_{N_{i+1}}$ (матрица $N_{i+1} \times N_i$), $i = 0, \ldots, L$,
- $N_0 = n,\ N_{L+1} = m$,
- f_i — **нелинейная** функция одной переменной $\mathbb{R} \to \mathbb{R}$, действующая “поэлементно”:

$$f_i\left((y_1, \ldots, y_{N_i})^T\right) = (f_i(y_1), \ldots, f_i(y_{N_i}))^T$$

Графическое представление нейронной сети

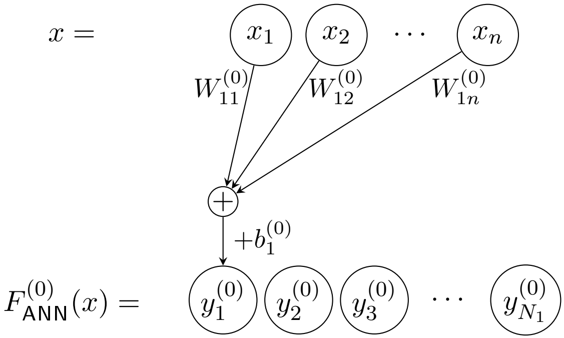
$$F^{(0)}_{\mathsf{ANN}}(x) = W^{(0)}x + b^{(0)}$$

$$x = \quad x_1 \quad x_2 \quad \cdots \quad x_n$$

$$F^{(0)}_{\mathsf{ANN}}(x) = \quad y^{(0)}_1 \quad y^{(0)}_2 \quad y^{(0)}_3 \quad \cdots \quad y^{(0)}_{N_1}$$

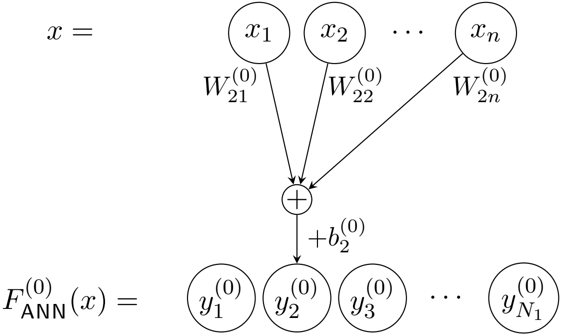
Графическое представление нейронной сети

$$F_{\mathsf{ANN}}^{(0)}(x) = W^{(0)}x + b^{(0)}$$



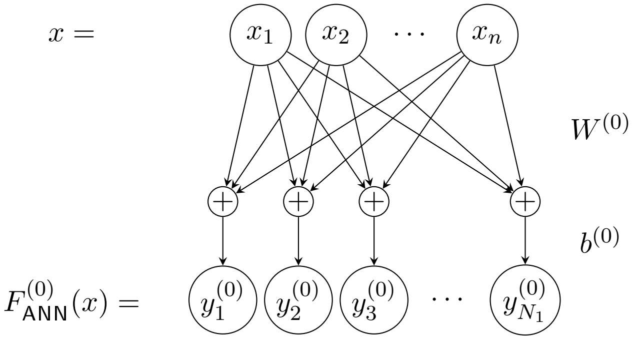
Графическое представление нейронной сети

$$F^{(0)}_{\mathsf{ANN}}(x) = W^{(0)}x + b^{(0)}$$



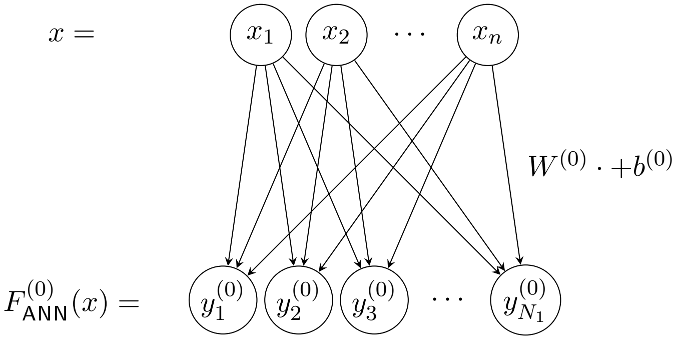
Графическое представление нейронной сети

$$F^{(0)}_{\mathsf{ANN}}(x) = W^{(0)}x + b^{(0)}$$



Графическое представление нейронной сети

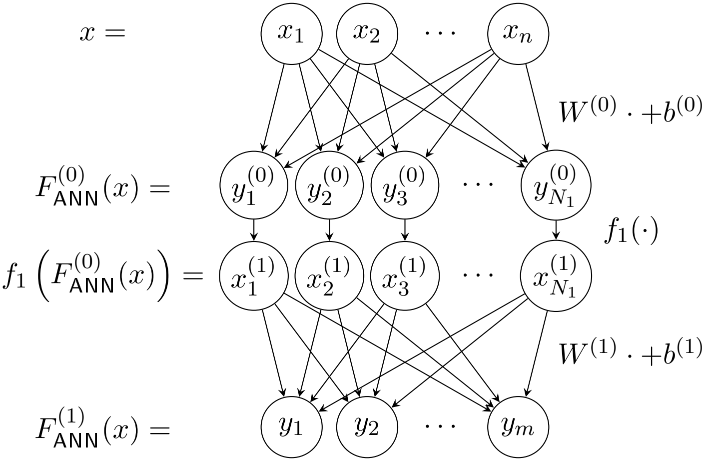
$$F^{(0)}_{\mathsf{ANN}}(x) = W^{(0)}x + b^{(0)}$$



Перцептрон Розенблатта: $L = 1$

F. Rosenblatt, “The perceptron. a perceiving and recognizing automaton (project para)”, Cornell Aeronautical Laboratory, Inc., Tech. Rep. 85-460-1, Jan. 1957

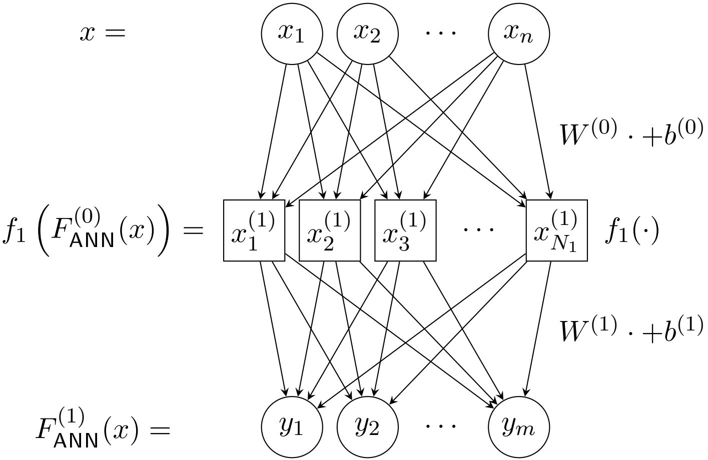
$$F^{(1)}_{\mathsf{ANN}}(x) = W^{(1)} f_1 \left(F^{(0)}_{\mathsf{ANN}}(x) \right) + b^{(1)}$$



Перцептрон Розенблатта: $L = 1$

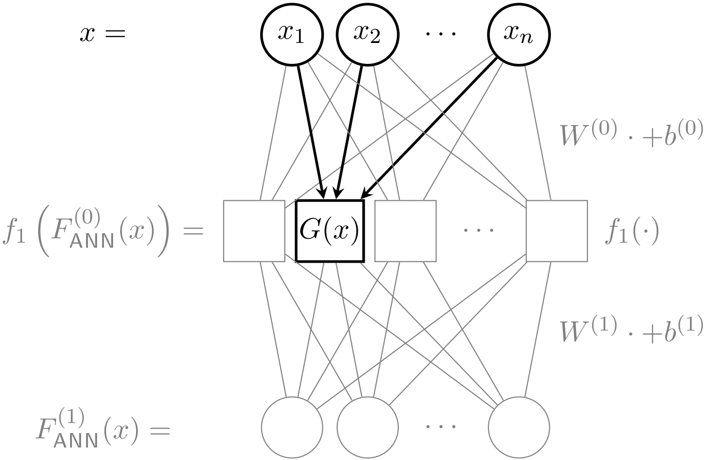
F. Rosenblatt, "The perceptron. a perceiving and recognizing automaton (project para)",
Cornell Aeronautical Laboratory, Inc., Tech. Rep. 85-460-1, Jan. 1957

$$F^{(1)}_{\mathsf{ANN}}(x) = W^{(1)} f_1 \left(F^{(0)}_{\mathsf{ANN}}(x) \right) + b^{(1)}$$

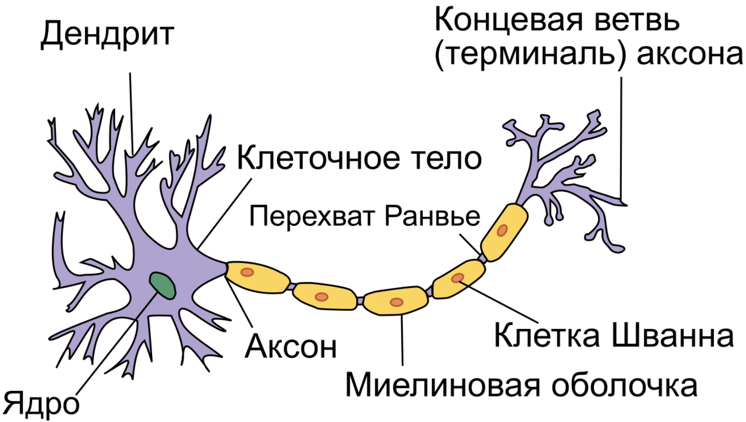


Искусственный нейрон

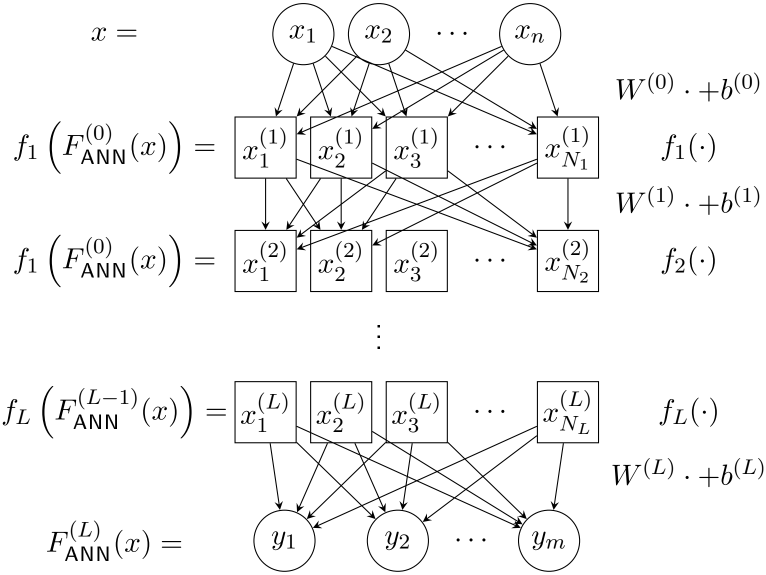
$$G(x) = f\left(wx + b\right) = f\left(\sum_{i=1}^{n} w_i x_i + b\right)$$



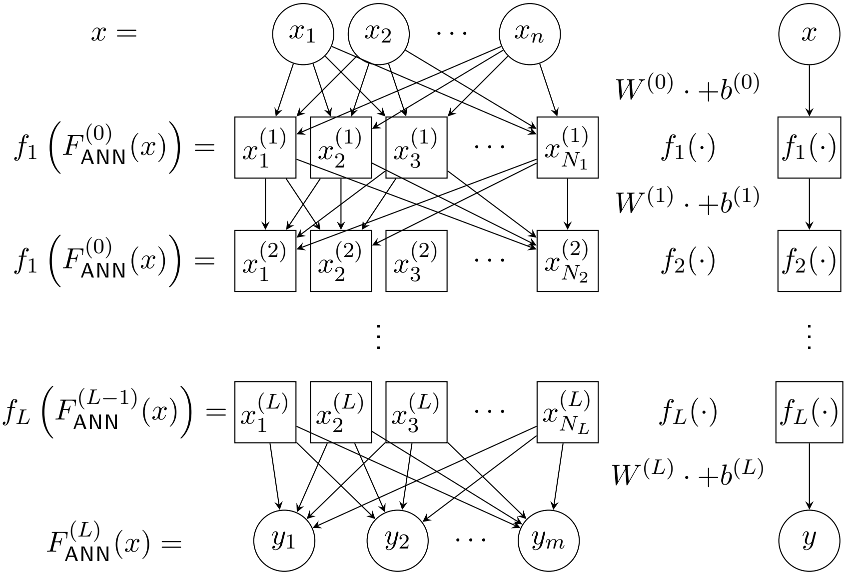
Естественный нейрон



Многослойный перцептрон: $L > 1$



Многослойный перцептрон: $L > 1$



Многослойный перцептрон: $L > 1$

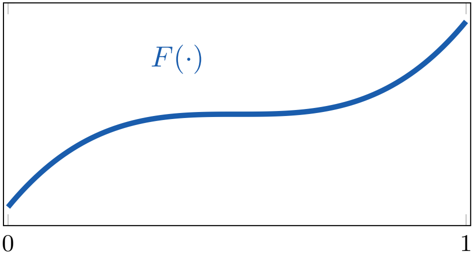
$$\begin{tikzcd}
x \arrow[d, "W^{(0)} \cdot + b^{(0)}"'] \\
\boxed{f_1(\cdot)} \arrow[d, "W^{(1)} \cdot + b^{(1)}"'] \\
\boxed{f_2(\cdot)} \\
\vdots \\
\boxed{f_L(\cdot)} \arrow[d, "W^{(L)} \cdot + b^{(L)}"'] \\
y
\end{tikzcd}$$

$$F_{\mathsf{ANN}}^{(L)}(x) =$$

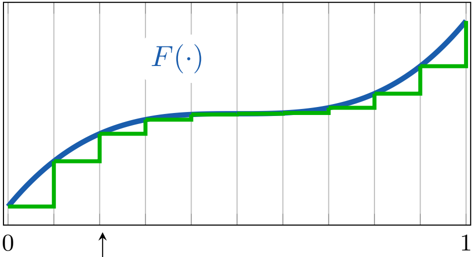
Аппроксимация функций многих переменных функциями меньшего числа переменных

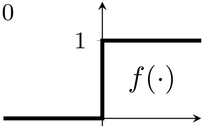
- 13-я проблема Гильберта (II Международный конгресс математиков, 8 августа 1900 г.). О представлении функций нескольких переменных в виде суперпозиций функций двух переменных. Решена Колмогоровым и Арнольдом в 1956-1957 г.
- Колмогоров. О представлении непрерывных ф-ций нескольких переменных в виде суперпозиций непрерывных функций одного переменного и сложения. 1957 г.
- Универсальная теорема аппроксимации. Три независимые работы в 1989 г.:
  - Цыбенко,
  - Фунахаши (Funahashi),
  - Хорник, Стинчкомб, Уайт (Hornik, Stinchcombe, White).
- 1998 г. Горбань. Обобщённая теорема аппроксимации. Аппроксимация с помощью линейных функций, суперпозиции и одной произвольной непрерывной **нелинейной** функции одной переменной.

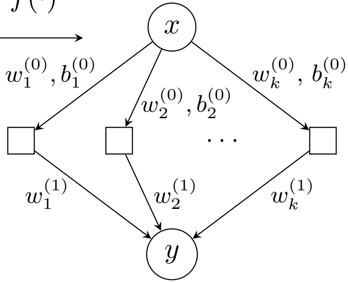
Принцип доказательства теоремы аппроксимации



Принцип доказательства теоремы аппроксимации







1. Кусочно-постоянная функция — перцептрон Розенблатта $[0, 1] \to \mathbb{R}$, имеющий k нейронов на скрытом слое.
2. $[0, 1]$ разбивается на k интервалов:

$$I_i = \Big[(i-1)/k, i/k\Big), \ i = 1, \dots, k$$

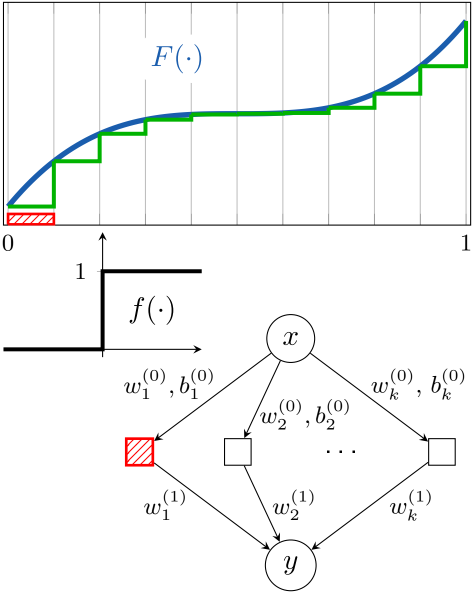
3. $w_i^{(0)} = 1, \ b_i^{(0)} = -(i-1)/k$

$$f\left(w_i^{(0)} x + b_i^{(0)}\right) = \begin{cases} 1, & x \in I_\alpha, \ \alpha \geqslant i \\ 0, & \text{иначе} \end{cases}$$

4. Имеем СЛАУ относительно $w_\alpha^{(1)}$

$$\sum_{\alpha=1}^{i} w_\alpha^{(1)} = F\big((i-1)/k\big), \ i = 1, \dots, k$$

Принцип доказательства теоремы аппроксимации



1. Кусочно-постоянная функция — перцептрон Розенблатта $[0,1] \to \mathbb{R}$, имеющий k нейронов на скрытом слое.
2. $[0,1]$ разбивается на k интервалов:

$$I_i = \Big[(i-1)/k, i/k\Big),\ i = 1, \dots, k$$

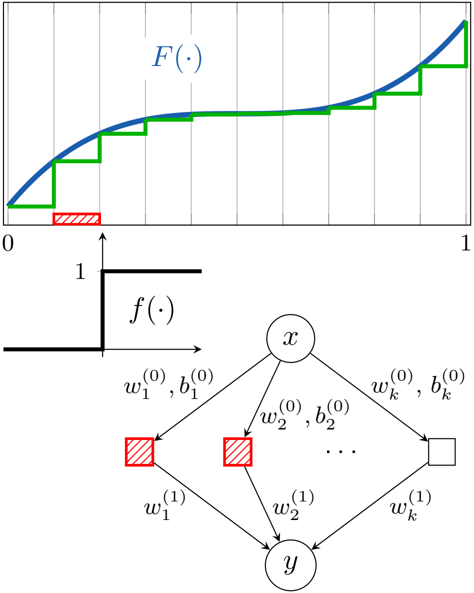
3. $w_i^{(0)} = 1,\ b_i^{(0)} = -(i-1)/k$

$$f\left(w_i^{(0)} x + b_i^{(0)}\right) = \begin{cases} 1, & x \in I_\alpha,\ \alpha \geqslant i \\ 0, & \text{иначе} \end{cases}$$

4. Имеем СЛАУ относительно $w_\alpha^{(1)}$

$$\sum_{\alpha=1}^{i} w_\alpha^{(1)} = F\big((i-1)/k\big),\ i = 1, \dots, k$$

Принцип доказательства теоремы аппроксимации



1. Кусочно-постоянная функция — перцептрон Розенблатта $[0,1] \to \mathbb{R}$, имеющий k нейронов на скрытом слое.
2. $[0,1]$ разбивается на k интервалов:

$$I_i = \Big[(i-1)/k, i/k\Big), \; i = 1, \ldots, k$$

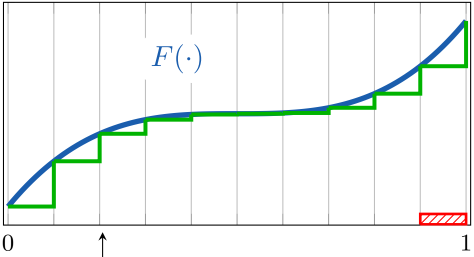
3. $w_i^{(0)} = 1, \; b_i^{(0)} = -(i-1)/k$

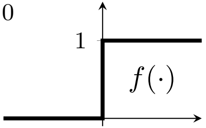
$$f\left(w_i^{(0)} x + b_i^{(0)}\right) = \begin{cases} 1, & x \in I_\alpha, \; \alpha \geqslant i \\ 0, & \text{иначе} \end{cases}$$

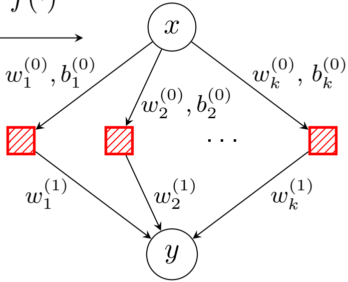
4. Имеем СЛАУ относительно $w_\alpha^{(1)}$

$$\sum_{\alpha=1}^{i} w_\alpha^{(1)} = F\big((i-1)/k\big), \; i = 1, \ldots, k$$

Принцип доказательства теоремы аппроксимации







1. Кусочно-постоянная функция — перцептрон Розенблатта $[0,1] \to \mathbb{R}$, имеющий k нейронов на скрытом слое.
2. $[0,1]$ разбивается на k интервалов:

$$I_i = \Big[(i-1)/k, i/k\Big), \; i = 1, \dots, k$$

3. $w_i^{(0)} = 1, \; b_i^{(0)} = -(i-1)/k$

$$f\left(w_i^{(0)} x + b_i^{(0)}\right) = \begin{cases} 1, & x \in I_\alpha, \; \alpha \geqslant i \\ 0, & \text{иначе} \end{cases}$$

4. Имеем СЛАУ относительно $w_\alpha^{(1)}$

$$\sum_{\alpha=1}^{i} w_\alpha^{(1)} = F\big((i-1)/k\big), \; i = 1, \dots, k$$

Тест

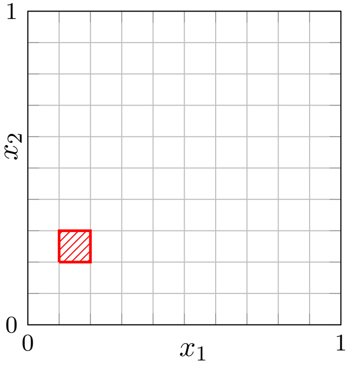
При прохождении теста укажите e-mail и фамилию, имя, отчество, которые вы указали при регистрации на курс. По этим данным будут суммироваться результаты всех ваших тестов в семестре. После ответа на тест вам на почту должно прийти письмо с вашими ответами.

https://docs.google.com/forms/d/e/1FAIpQLSfdtetaaUntxVZ_yBdXQHhd9SvLrFmB5FWfT7bfVQcBtHW-cg/viewform

https://neurofuzzy.phys.msu.ru/redmine/projects/ai_in_natural_science/news

Принцип доказательства теоремы аппроксимации

Двумерный случай



- $[0,1]\times[0,1]$ разбивается на k^2 квадратов:

$$I_{ij} = \left[\frac{i-1}{k}, \frac{i}{k}\right) \times \left[\frac{j-1}{k}, \frac{j}{k}\right), \quad i,\, j = 1, \ldots, k$$

- $f\Big(kx_1 - (i-1) + x_2 - (j-1)/k\Big) =$

$$= \begin{cases} 1, & (x_1, x_2) \in I_{\alpha\beta},\ \alpha \geqslant i \text{ и } \beta \geqslant j \\ 0, & \text{иначе} \end{cases}$$

n-мерный случай

- Количество нейронов на скрытом слое — k^n.
- Для изображения n — это количество пикселей ($\sim 10^6$ — мегапиксель). Следовательно, необходимо $\sim 2^{10^6}$ **нейронов**, что больше, чем количество атомов во вселенной.