

Онлайн-курс

# ВВЕДЕНИЕ В ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ

## 4. Искусственные нейронные сети

Автор: Шевляков Артём Николаевич

# Биологические нейроны

# Биологический нейрон

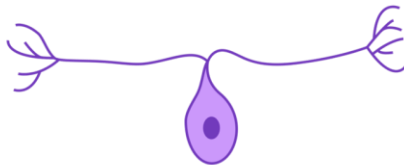
У него есть отростки (аксоны), которые ведут в другие нейроны. Из этих нейронов может приходиться нервный импульс. Нейрон может генерировать импульс и передавать его соседям.



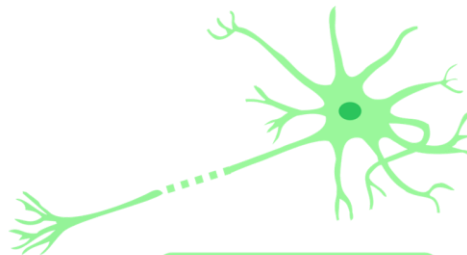
Биполярный



Униполярный



Псевдоуниполярный



Мультиполярный

# Биологический нейрон

**Важно: соседи нейрона не равноправны.**

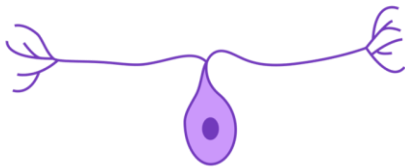
Важность соседа выражается в толщине отростка (импульс от важного соседа будет более сильным).



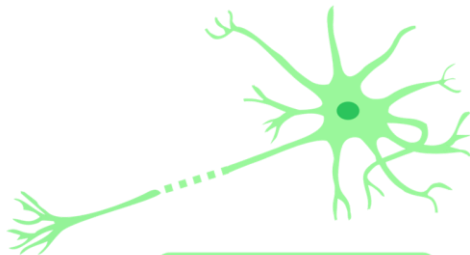
Биполярный



Униполярный



Псевдоуниполярный



Мультиполярный

# Биологический нейрон

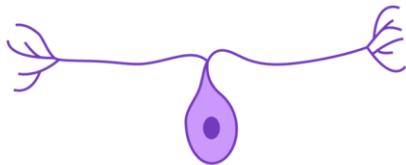
Каждый нейрон может находиться в двух состояниях: возбужденном и невозбужденном. Если в нейрон пришел слишком большой сигнал, то нейрон переходит в возбужденное состояние.



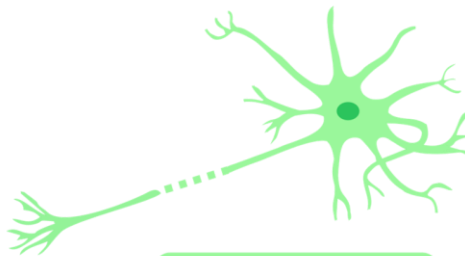
Биполярный



Униполярный



Псевдоуниполярный



Мультиполярный

# Биологический нейрон

**Нейроны могут образовывать цикл.**

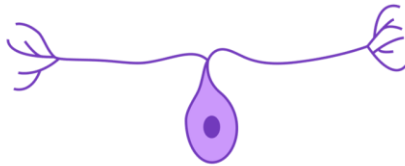
В этом случае импульс, испущенный из нейрона, может (после всех преобразований в нейронах цикла) снова вернуться в него.



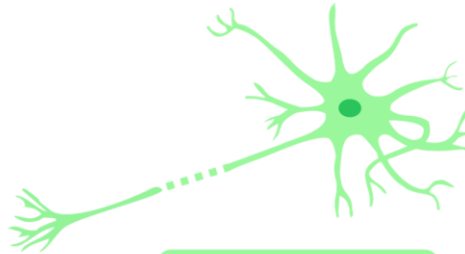
Биполярный



Униполярный



Псевдоуниполярный



Мультиполярный

Онлайн-курс

# ВВЕДЕНИЕ В ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ

## 4. Искусственные нейронные сети

# Искусственный нейрон



# А как представить нейрон в виде матем. объекта?

- У него должны быть входные и выходные значения.
- У него должна быть указана важность (сила) каждой связи с соседним нейроном.
- Он должен по достаточно простой формуле обрабатывать входные значения и передавать результат вычислений дальше.
- Выходное значение нейрона должно моделировать возбуждение нейрона.

# В общем, нужно записать такую штуку

У искусственного нейрона (ИН) каждый вход имеет свой **вес**  $w_i$ .

Входные значения  $x_i$  приходят либо из внешней среды либо из других нейронов.

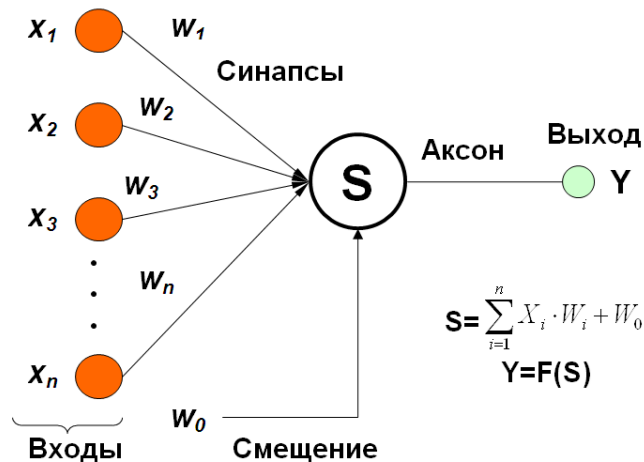
Важность каждого входа выражается весом связи  $w_i$ .

ИН считает взвешенную сумму входов, прибавляет **смещение**  $w_0$ .

Смысл смещения: **порог возбуждения**.

Далее к результату применяется **функции активации** (ФА).

Если результирующее значение достаточно велико, то считается, что нейрон возбудился.

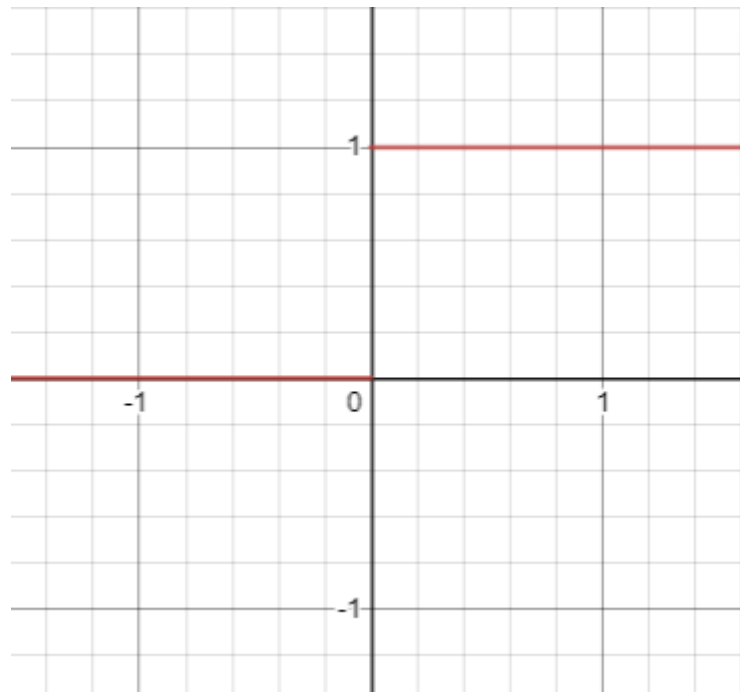


# Какую ФА взять?

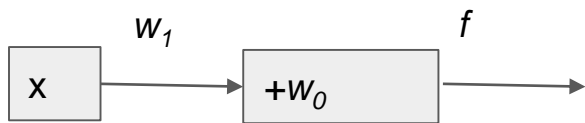
Если строго следовать биологии, мы должны взять ступенчатую функцию

$$step(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

она моделирует  
переход нейрона в  
возбужденное состояние



# Пример работы искусственного нейрона



Если вход нейрона  $x=1$ , его вес  $w_1=2$ , смещение  $w_0=-3$ , ФА  $f(x)=\text{step}(x)$ , то на выходе будет значение  $\text{step}(2*1-3)=0$ . То есть нейрон не возбуждается.

А вот если бы смещение было бы  $w_0=-1$ , возбуждение ему обеспечено:  $\text{step}(2*1-1)=1$ .

**Отсюда смысл смещения:** мы настраиваем границу входного сигнала, при превышении которой происходит возбуждение.

# Недостатки функции-ступеньки step

- Она разрывна.
- Её производная равна 0 во всех точках своей области определения.
- Следовательно, бесполезно использовать ГС при минимизации выражений, содержащих функцию-ступеньку.

**Но позвольте, зачем нужно минимизировать ИН?**

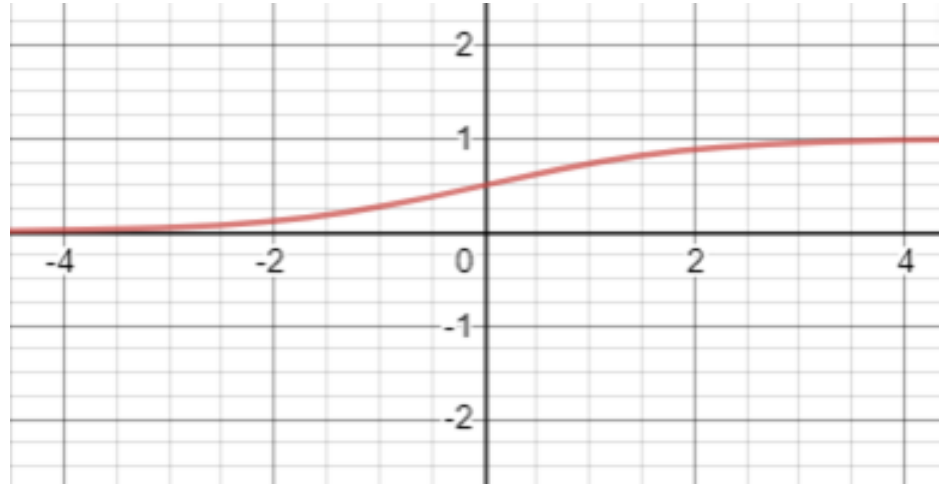
Это нужно для поиска их оптимальных параметров (весов и смещений), необходимых для построения модели искусственного интеллекта.

Ну что ж. **Будем искать замену ступеньке.** Сигмоида подойдет?

# Сигмоида как функция активации

Сигмоида является гладкой аппроксимацией ступеньки. В исторически первых НС применялась именно она.

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



# Недостатки сигмоиды

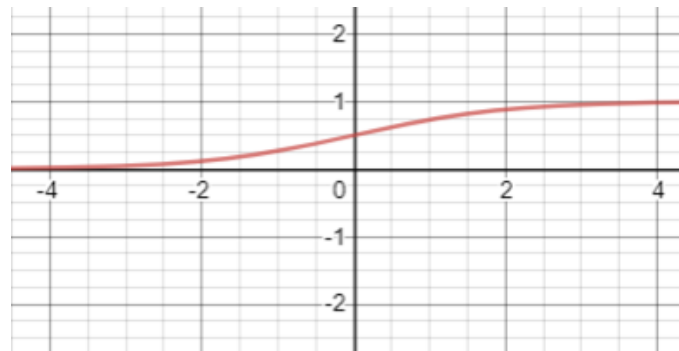
В предыдущих лекциях мы говорили о том, что сигмоида не безупречна.

Когда у вас будет нейронная сеть, в которой несколько нейронов соединены последовательно, то возникнет суперпозиция сигмоид

$$\sigma(\sigma(\sigma(x)))$$

А у такой функции градиент практически нулевой.

Есть ли еще другие популярные ФА?



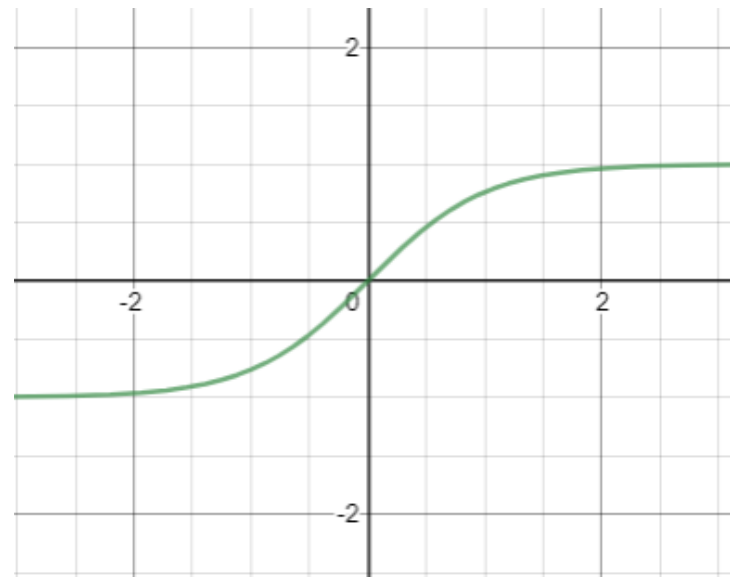
# Гиперболический тангенс

$$th(x) = \frac{sh(x)}{ch(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

он как бы тоже аппроксимирует ступеньку.

Такая ФА реально применяется в больших НС (LeNet)

А можно брать «не лежащие»  
функции активации?





# ReLU - функция

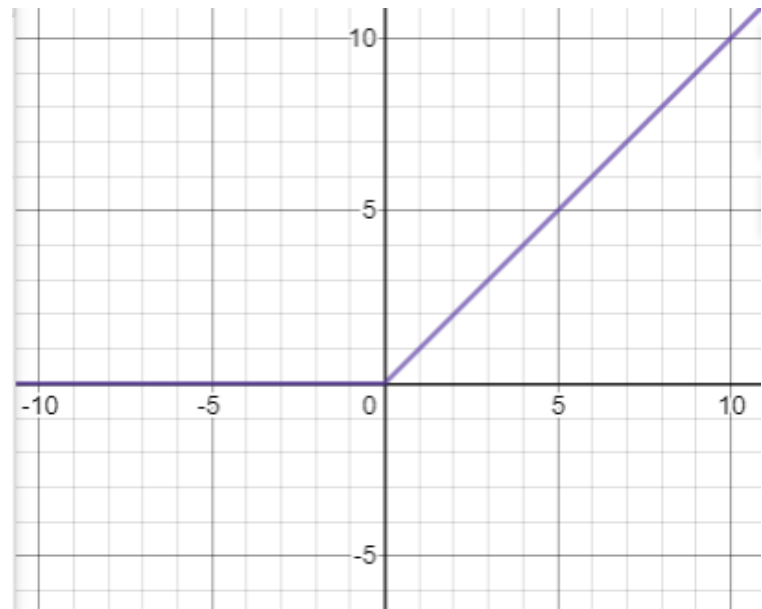
ReLU=Rectified linear unit

**Достоинства:** простота.

**Недоумение:** она совсем не похожа на ступеньку.

Практика показала, что часто оправдано моделировать ИН не двумя состояниями (возбуждился или нет), а числовой «шкалой возбуждения».

ReLU для этого идеально подходит.



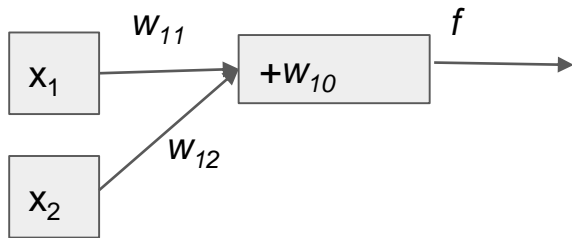
# Пример работы искусственного нейрона



Если вход нейрона  $x=1$ , его вес  $w_1=2$ , смещение  $w_0=-3$ , ФА  $f(x)=Relu(x)$ , то на выходе будет значение  $Relu(2*1-3)=0$ . То есть нейрон не возбуждается.

А вот если бы смещение было бы  $w_0=1$ , то выход нейрона будет  $Relu(2*1+1)=3$ .

# Несколько входов искусственного нейрона



У ИН может быть несколько входов (каждый со своим весом). Входы умножаются на веса и суммируются.

**Например,** для  $x_1=1$ ,  $x_2=-1$ ,  $w_{11}=-2$ ,  $w_{12}=3$ ,  $w_{10}=10$ ,  $f=Relu$

получаем  $Relu(1*(-2)+(-1)*3+10)=Relu(-5+10)=5$ .

Онлайн-курс

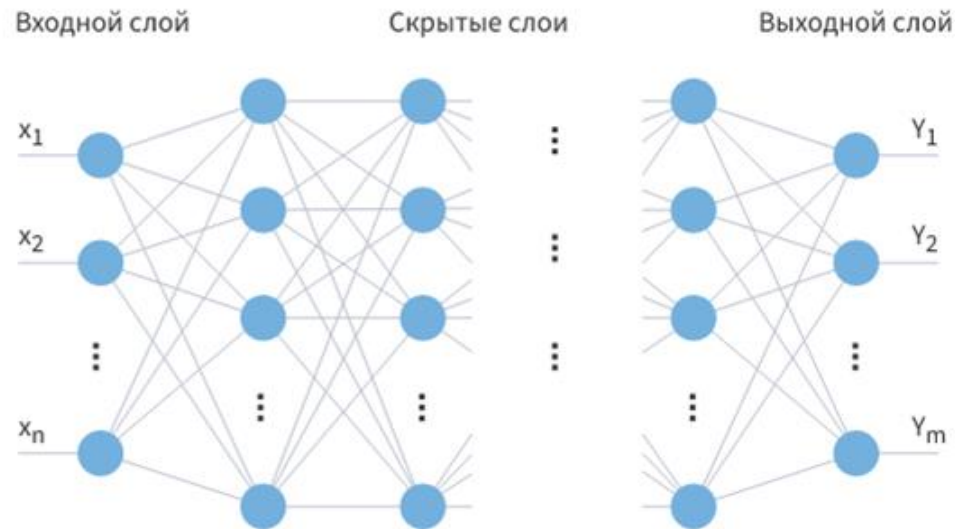
# ВВЕДЕНИЕ В ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ

## 4. Искусственные нейронные сети

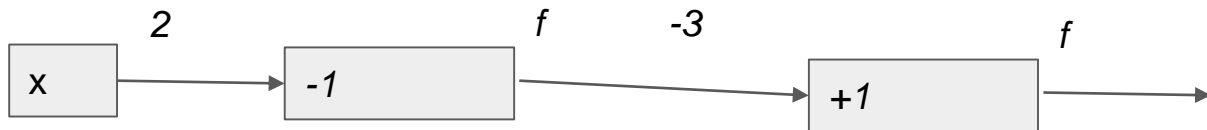
# Нейронные сети. Общие сведения

# Что можно сделать с искусственными нейронами?

Их можно соединить в **нейронную сеть** (НС) так, что выход одного нейрона является входом другого нейрона.



# Пара нейронов



Пусть все функции активации Relu.

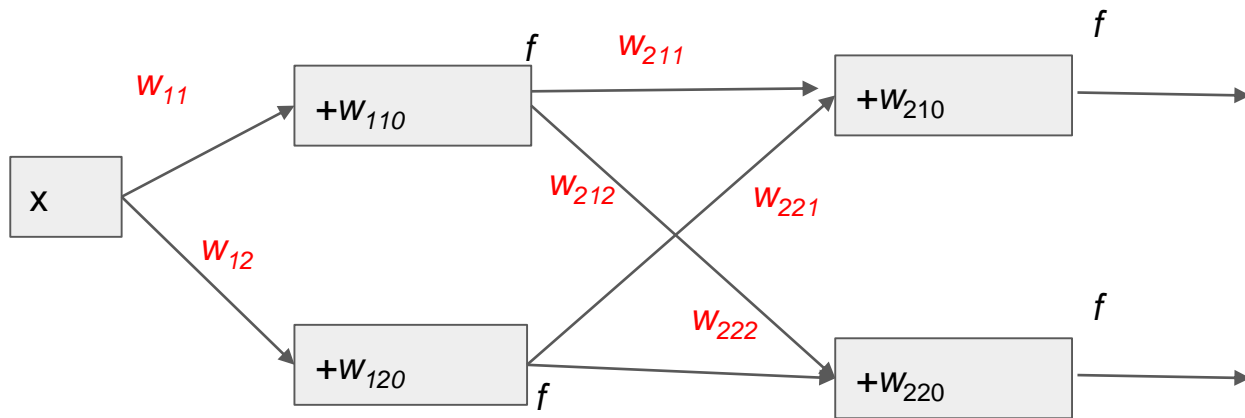
Тогда для входа  $x=1$  НС выдаст ответ:

$$\text{Relu}((-3) * \text{Relu}(2 * 1 - 1) + 1) = \text{Relu}(-2) = 0$$

# Веса нейронной сети

В НС есть два типа весов:

**веса-связи** (веса связи между двумя нейронами) и **веса-смещения**.

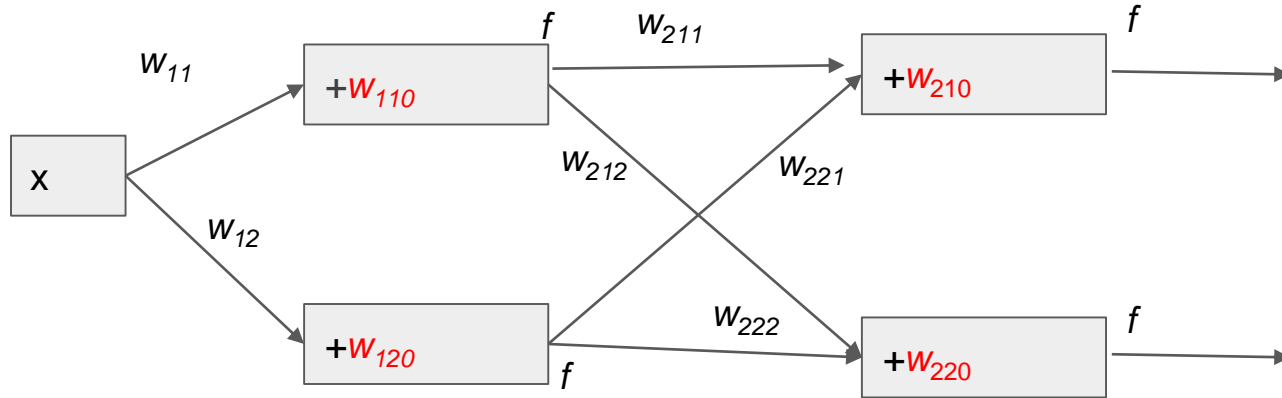




# Веса нейронной сети

В НС есть два типа весов:

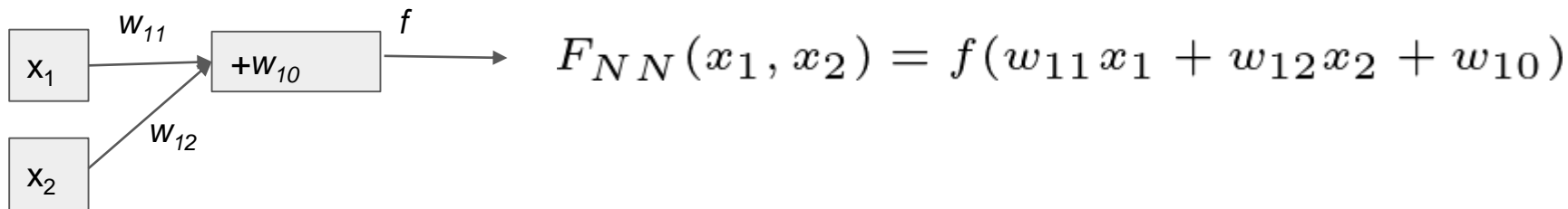
**веса-связи** (веса связи между двумя нейронами) и **веса-смещения**.



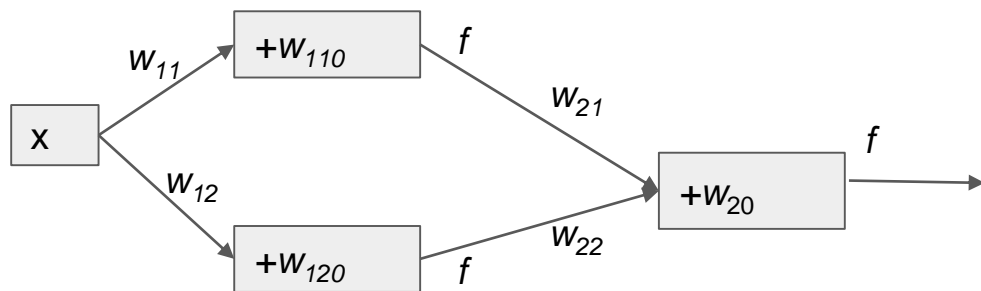
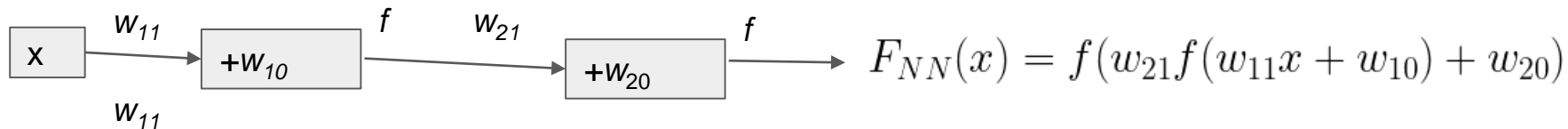
# Функция сети

С НС можно связать функцию  $F_{NN}(x)$ , которая преобразует вход сети в выход.

Например:



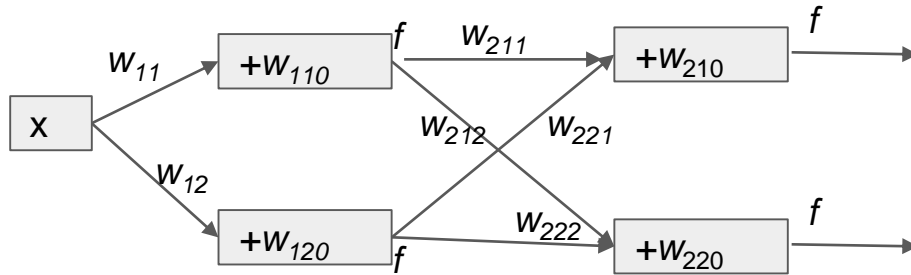
# Простейшие сети и их функции



$$F_{NN}(x) = f(f(w_{11}x + w_{110})w_{21} + f(w_{12}x + w_{120})w_{22} + w_{20})$$

# НС с несколькими выходами

Если у НС несколько выходов (а такое бывает), то ее функция  $F_{NN}(x)$  является вектором.



$$F_{NN}(x) = (F_1, F_2),$$

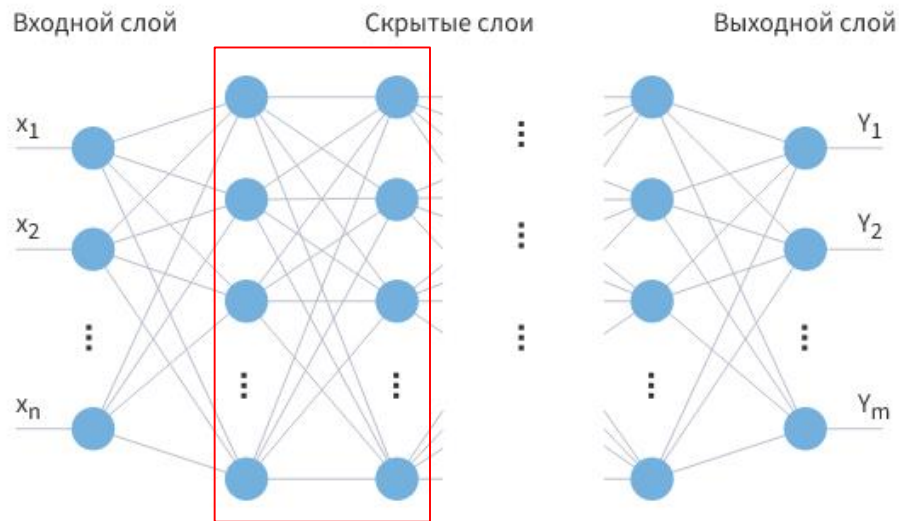
$$F_1 = f(w_{211} f(w_{11}x + w_{110}) + w_{221} f(w_{12}x + w_{120}) + w_{210}),$$

$$F_2 = f(w_{212} f(w_{11}x + w_{110}) + w_{222} f(w_{12}x + w_{120}) + w_{220})$$

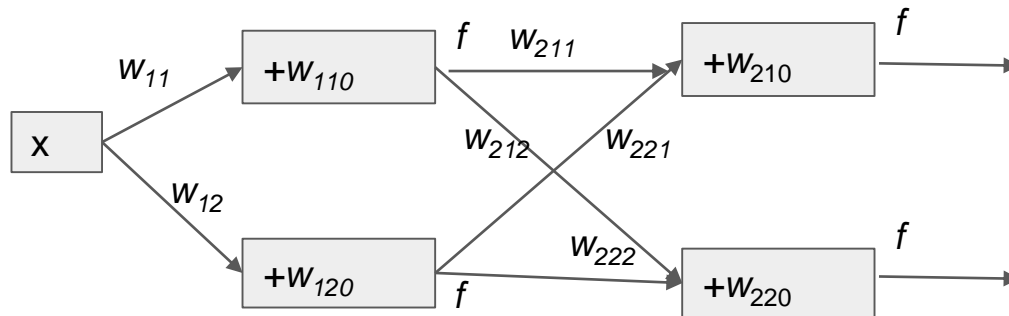
# Полносвязные НС

Если каждый нейрон некоторого слоя НС **соединен со всеми** нейронами следующего слоя, то такая пара слоев называется **полносвязной**.

В паре полносвязных слоев может быть разное число нейронов.



# Число весов в полносвязных слоях



Если два слоя полносвязны, то число весов в связях между ними равно  $n_1 n_2$ , где  $n_1, n_2$  - количество нейронов в этих слоях.

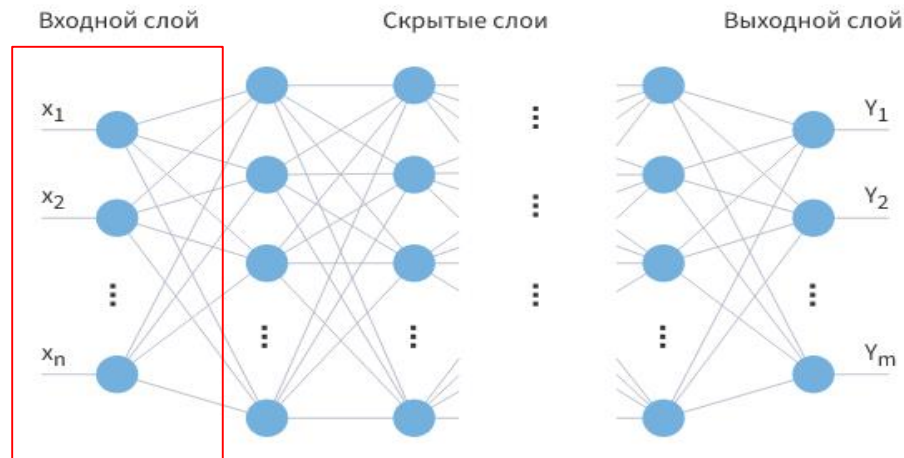
Нужно не забыть, что еще есть веса-смещения. Их число совпадает с общим числом нейронов в НС (кроме нейронов входного слоя).

# Входной слой

Входной слой – особый. Там лежат входные данные. **Размер входного слоя совпадает с размерностью данных.**

Если в НС подаются числа, то размерность входного слоя равна 1.

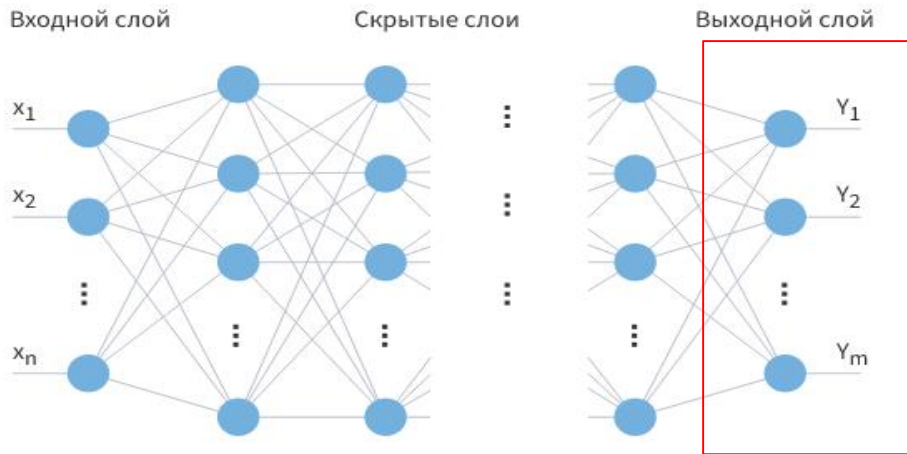
Если в НС подаются пары чисел, то размерность входного слоя равна 2  
и.т.д.



# Выходной слой

Выходной слой – тоже особый. **То, что выходит с этого слоя, считается ответом все НС.** Размер выходного слоя может быть разным – в зависимости от типа ответа.

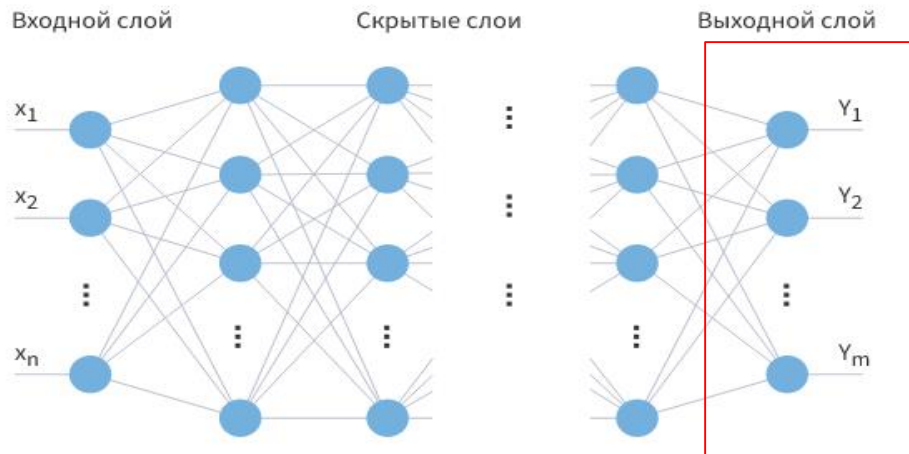
Если ответ сети – единственное число, то на выходном слое будет 1 нейрон. Но это не обязательно...





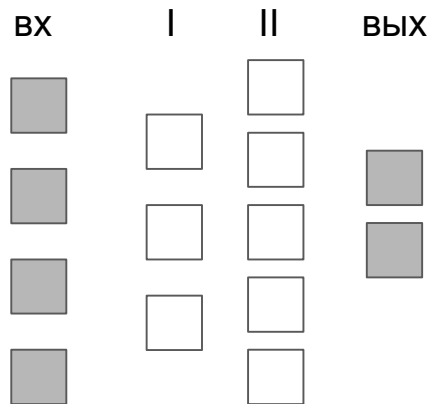
# Зачем у НС несколько выходов

Несколько выходов у НС оправданы, когда НС выдает массив чисел. Например, НС может выдать список вероятностей, что поданная на вход фотография содержит изображение собаки, кошки, автомобиля соответственно.



# Задача

Найти общее количество весов НС (веса связи между нейронами и смещения), если НС состоит из входного слоя размерности 4, двух внутренних слоев размеров 3 и 5 соответственно, и выходного слоя размерности 2. Все слои в НС полносвязны.

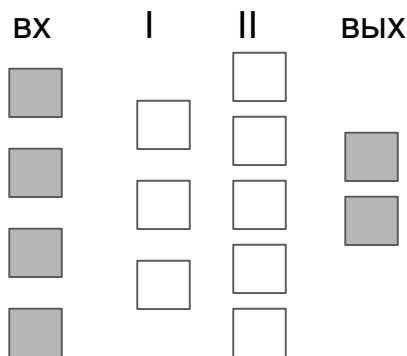


# Задача

Веса-смещения есть у каждого нейрона, кроме нейронов входного слоя. Итого: 10 весов смещений.

Число весов-связей между парой слоев равно произведению числа нейронов в этих слоях.

Будет  $4 \times 3 = 12$  весов-связей между входным слоем и первым внутренним слоем.



Будет  $3 \times 5 = 15$  весов-связей между первым и вторым внутренним слоем.

Будет  $5 \times 2 = 10$  весов-связей между вторым внутренним слоем и выходным слоем.

**Всего весов:**  $10 + 12 + 15 + 10 = 47$

# Зачем считать количество весов в НС?

Во время тренировки НС ищутся оптимальные значения весов сети. Чем больше число весов НС, тем больше...

- ... время тренировки НС;
- ... требуется данных для НС;
- ... (сюрприз!) вероятность того, что НС обнаружит ложную закономерность (переобучение, overfitting).

Онлайн-курс

# ВВЕДЕНИЕ В ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ

## 4. Искусственные нейронные сети

# **Различия между биологическими и искусственными нейронными сетями**

# Похож ли искусственный нейрон на биологический?

**Чем человеческие мозги лучше искусственных?**

(Пока такой вопрос не считается расизмом)

Сходство в строении, бесспорно, есть. Но есть и принципиальные различия.



# Различие №1

Сигнал в искусственной НС распространяется в одном направлении, то есть **искусственные нейроны не могут образовывать циклы.**

А что из этого следует?



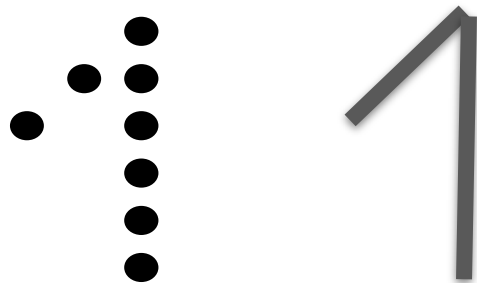
# Различие №1

Например, искусственная НС не подвержена галлюцинациям.

Но это не самое важное.

**Искусственная нейронная сеть не может хорошо достраивать изображения.**

Например, эти изображения для искусственной НС принципиально различны. И сделать, так, чтобы НС начала в них видеть сходство – весьма нетривиальная задача.



# Различие №2

Практика показала, что для искусственных НС два состояния нейрона (возбужденный – невозбужденный) недостаточно.

**Поэтому de facto была принята непрерывная «шкала возбуждения» и теперь наиболее популярна такая функция активации (Relu).**

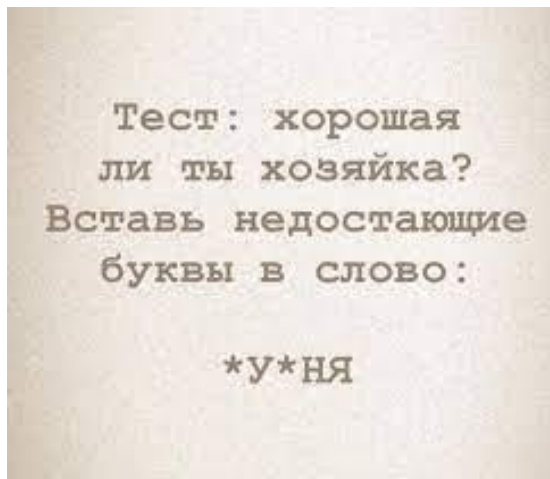
# Различие №3

Для настройки весов искусственной НС используется **алгоритм обратного распространения**.

Никаких аналогов такого алгоритма в биологических НС нет!

# Различие №4

Чтобы найти истинную закономерность в данных, человеческому мозгу требуется на порядок меньше данных.



Онлайн-курс

# ВВЕДЕНИЕ В ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ

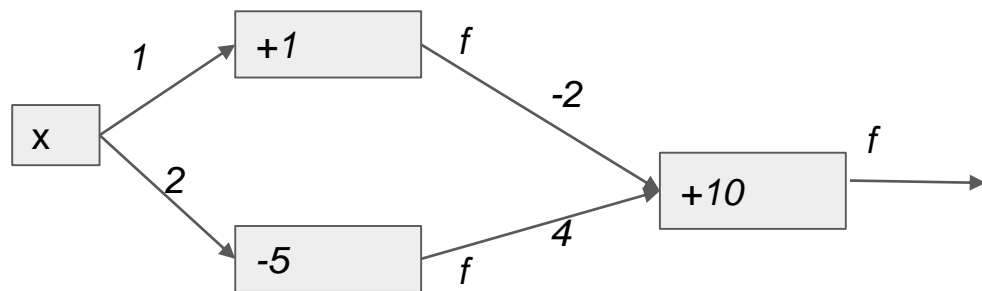
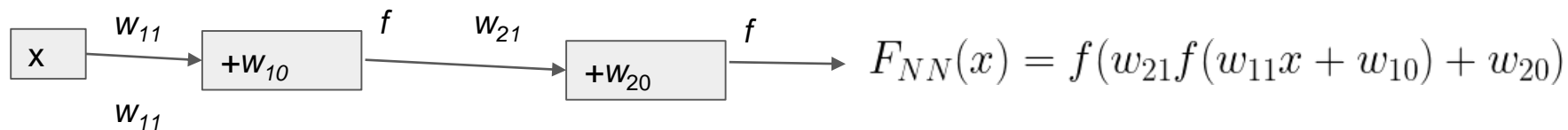
## 4. Искусственные нейронные сети

# Выводы

# Выводы:

- Мы обсудили строение биологического нейрона.
- Было введено определение искусственного нейрона.
- Мы рассмотрели различные типы функций активаций.
- Научились соединять несколько искусственных нейронов в нейронную сеть.
- Сравнили биологические и искусственные нейронные сети между собой.

# Простейшие сети и их функции



$$F_{NN}(x) = f(f(w_{11}x + w_{110})w_{21} + f(w_{12}x + w_{120})w_{22} + w_{20})$$