

# Тятя! Тятя! Наши сети притащили мертвеца!

## Задача 1

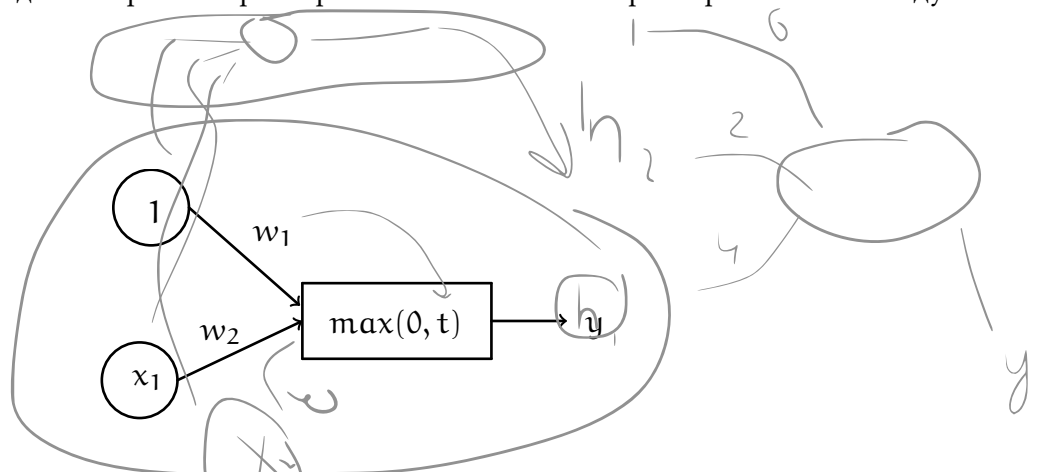
Маша услышала про машин лёрнинг и решила, что они и есть та самая Маша, которой этот лёрнинг принадлежит. Теперь она собрала два наблюдения:  $x_1 = 1, x_2 = 2, y_1 = 2, y_2 = 3$  и собирается обучить линейную регрессию  $y = \beta \cdot x$ . Она собирается сделать это тремя способами, и ей нужна ваша помощь!

1. Получить теоретическую оценку методом наименьших квадратов.
2. Методом градиентного спуска. Она собирается в качестве скорости обучения взять  $\eta = 0.1$ . В качестве стартовой точки она хочет использовать  $\beta_0 = 0$ . Обучение заканчивается после первого шага.
3. Методом стохастического градиентного спуска. Все параметры берутся такими же как в предыдущем пункте. Делается два шага. Сначала с первым наблюдением, потом со вторым.

## Задача 2

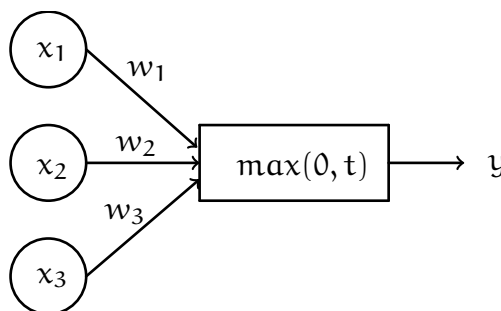
Парни очень любят Олю, а Оля любит собирать перцептроны и думать по вечерам о их весах и функциях активации. Сегодня она решила разобрать свои залежи из перцептронов и как следует упорядочить их.

- Для перцептрона



нужно подобрать веса так, чтобы он превращал  $x_1 = 0$  в  $y = 1$ , а  $x_1 = 1$  в  $y = 0$ .

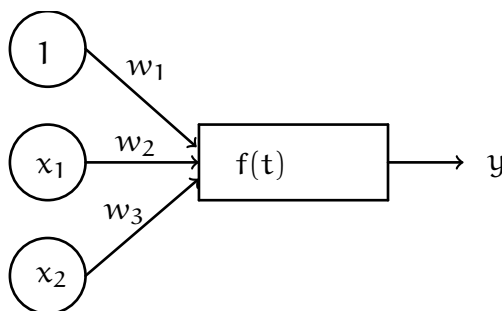
- Для перцепторона



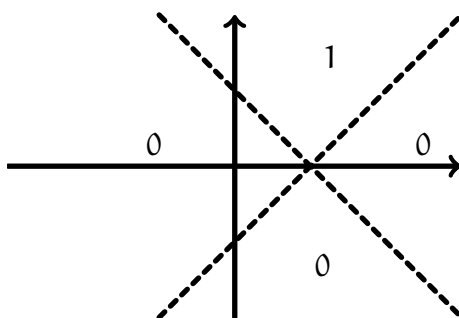
Оля хочет по наблюдениям  $x$  подобрать такие веса  $w_i$ , чтобы на выходе получились  $y$ .

$x_1$	$x_2$	$x_3$	$y$
1	1	2	0.5
1	-1	1	0

- У Оли есть несколько вот таких перцептронов с неизвестной функцией активации (надо самому выбирать):



На плоскости проведены две прямые  $x_1 + x_2 = 1$  и  $x_1 - x_2 = 1$ .



Оле нужно собрать нейросетку, которая будет классифицировать объекты с плоскости так, как показано на картинке.

### Задача 3

Попробуйте с помощью нейросеток с минимально возможным числом нейронов описать логические функции, заданные следующими таблицами истинности:

$x_1$	$x_2$	$x_1 \cap x_2$
1	1	1
1	0	0
0	1	0
0	0	0

$x_1$	$x_2$	$x_1 \cup x_2$
1	1	1
1	0	1
0	1	1
0	0	0

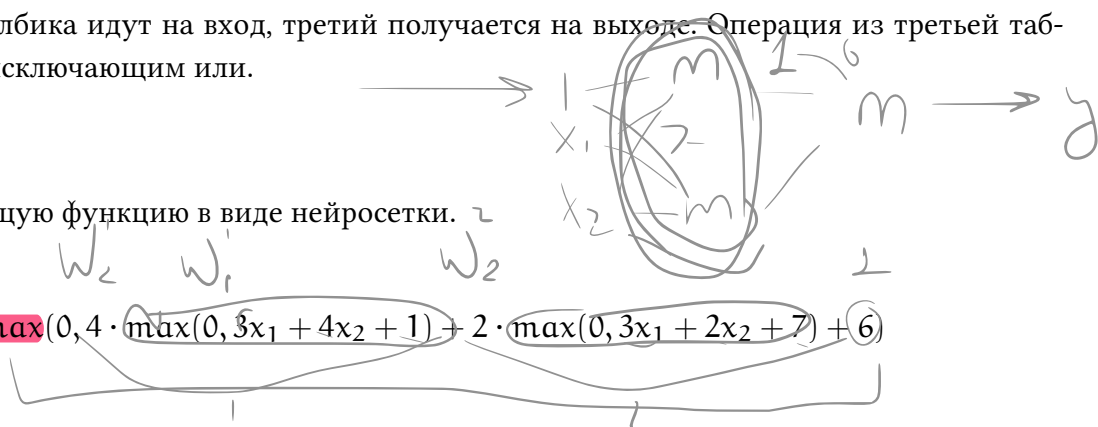
$x_1$	$x_2$	$x_1 \text{ XOR } x_2$
1	1	0
1	0	1
0	1	1
0	0	0

Первые два столбика идут на вход, третий получается на выходе. Операция из третьей таблицы называется исключающим или.

### Задача 4

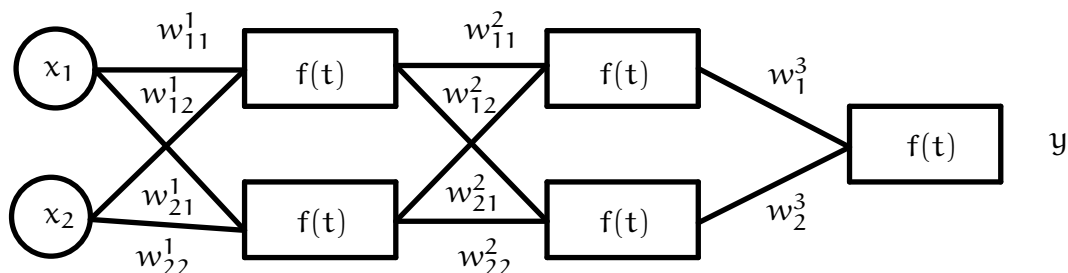
Нарисуйте следующую функцию в виде нейросетки.

$$y = \max(0, 4 \cdot \max(0, 3x_1 + 4x_2 + 1)) + 2 \cdot \max(0, 3x_1 + 2x_2 + 7) + 6$$



## Задача 5

Дана нейросетка:



1. Перепишите её как сложную функцию.
2. Запишите эту функцию в матричном виде.
3. Предположим, что  $L(W_1, W_2, W_3) = \frac{1}{2} \cdot (y - \hat{y})^2$  — функция потерь, где  $W_i$  — веса  $i$ -го слоя. Найдите производную функции  $L$  по всем весам  $W_i$ .

## Задача 6

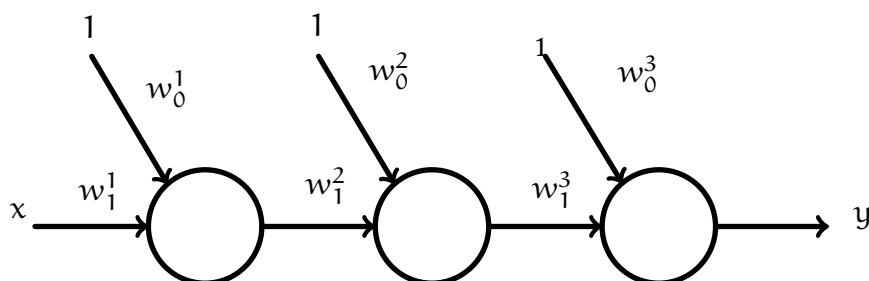
Изобразите для функции  $f(x, y) = x^2 + xy + (x + y)^2$  граф вычислений. Найдите производные всех выходов по всем входам. Опираясь на граф выпишите частные производные функции  $f$ .

## Задача 7

Функция  $f(t) = \frac{e^t}{1+e^t}$  называется сигмOIDом. Покажите, что  $f'(t) = f(t)(1 - f(t))$ .

## Задача 8

Как-то раз Вовочка решал задачу классификации. С тех пор у него в кармане завалялась нейросеть:

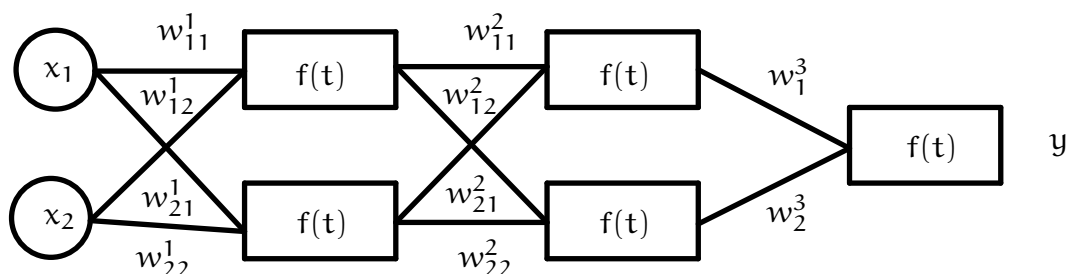


В качестве функции активации используется сигмOID:  $f(t) = \frac{e^t}{1+e^t}$ . Есть два наблюдения:  $x_1 = 1, x_2 = 5, y_1 = 1, y_2 = 0$ . Скорость обучения  $\gamma = 1$ . В качестве инициализации взяты нулевые веса. Как это обычно бывает, Вовочка обнаружил её в своих штанах после стирки и

очень обрадовался. Теперь он собирается сделать два шага стохастического градиентного спуска, используя алгоритм обратного распространения ошибки. Помогите ему.

### Задача 9

Дана нейросетка:



Для квадратичной функции ошибки  $MSE(W_1, W_2, W_3) = (y - \hat{y})^2$  выпишите все производные в том виде, в котором их было бы удобно использовать для алгоритма обратного распространения ошибки.

### Задача 10

Та, кому принадлежит машин лёрнинг собирается обучить нейронную сеть для решения задачи регрессии. На вход в ней идёт 12 переменных, в сетке есть 3 скрытых слоя. В первом слое 300 нейронов, во втором 200, в третьем 100.

- Сколько параметров предстоит оценить Маше? Сколько наблюдений вы бы на её месте использовали?
- Что Маша должна сделать с внешним слоем, если она собирается решать задачу классификации на два класса и получать на выходе вероятность принадлежности к первому классу?
- Что делать Маше, если она хочет решать задачу классификации на  $K$  классов?