

Тятя! Тятя! Наши сети притащили мертвеца!

Вышка
осень 2020

Задачи для второй посиделки

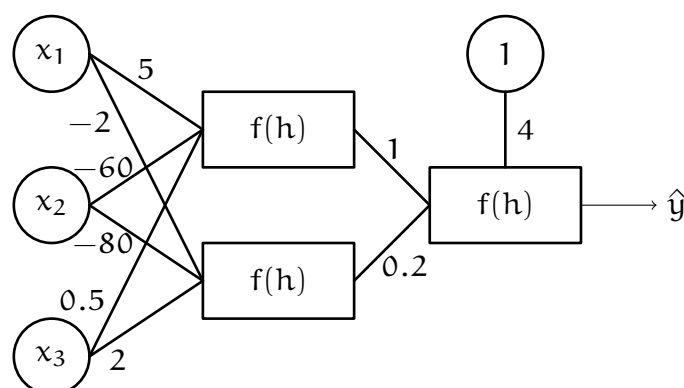
Задача 1

Маша слышала 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.
4. Сами на листочке за чашкой чая сделайте всё то же самое для оставшихся 50 видов градиентного спуска.

Задача 2

Добродум хозяин кофейни на Тверской. Он хочет понять насколько сильно будет заполнена кофейня в следующие выходные. Для этого по старым данным он обучил нейросетку. На вход она принимает три фактора: температуру за окном, x_1 , факт наличия на Тверской митинга, x_2 и пол баристы на смене, x_3 . В качестве функции активации Добродум использует ReLU.



1. В эти выходные за барной стойкой стоит Агнесса. Митинга не предвидится, температура будет в районе 20 градусов. Сколько человек придёт в кофейню к Добродуму?
2. На самом деле каждая нейросетка — это просто-напросто какая-то нелинейная сложная функция. Запишите нейросеть Добродума в виде функции.

Задача 3

Теперь в обратную сторону. Пусть у нас есть вот такая функция.

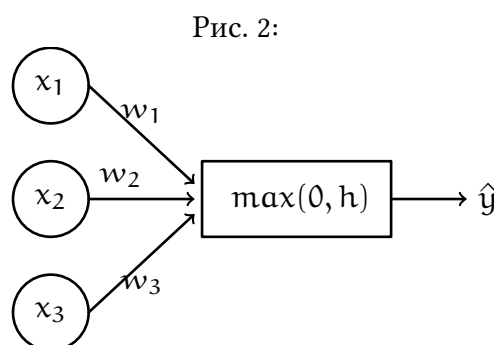
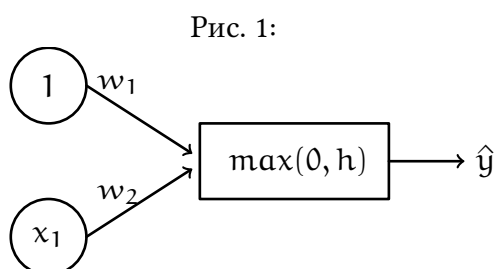
$$y = \max(0, 4 \cdot \max(0, 3x_1 + 4x_2 + 1) + 2 \cdot \max(0, 3x_1 + 2x_2 + 7) + 6)$$

Нарисуйте эту функцию в виде нейросетки.

Задача 4

Парни очень любят Машу, а Маша с недавних пор любит собирать персептроны и думать по вечерам об их весах и функциях активации. Сегодня она решила разобрать свои залежи из персептронов и как следует упорядочить их.

1. В ящике стола Маша нашла персептрон с картинки 1 Маша хочет подобрать веса так, чтобы он реализовывал логическое отрицание, то есть превращал $x_1 = 0$ в $y = 1$, а $x_1 = 1$ в $y = 0$.



2. В тумбочке среди носков Маша нашла персептрон, с картинки 2, Маша хочет подобрать такие веса w_i , чтобы персептрон превращал x из таблички в соответствующие y :

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

3. Оказывается, что в ванной всё это время валялась куча персептронов с картинки 3 с неизвестной функцией активации (надо самому выбирать).

Рис. 3:

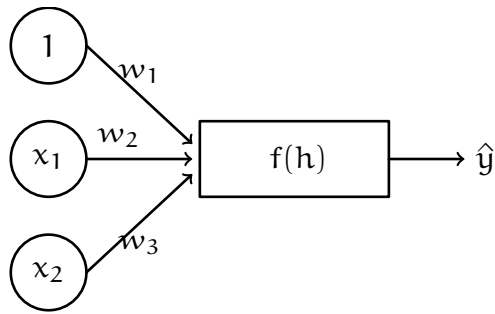
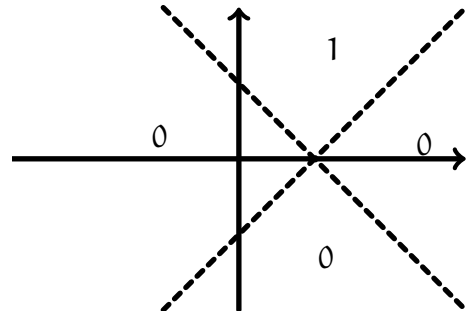


Рис. 4:



Маша провела на плоскости две прямые: $x_1 + x_2 = 1$ и $x_1 - x_2 = 1$. Она хочет собрать из персептронов нейросетку, которая будет классифицировать объекты с плоскости так, как показано на картинке 4.

Задача 5

Маша задумалась о том, можно ли с помощью нейронной сетки с одним скрытым слоем и ступенчатой функцией активации решить абсолютно любую задачу классификации на два класса со сколь угодно большой точностью? Ей кажется, что да. Как это можно сделать?

Задача 6

Доказать, что с помощью однослойной нейронной сетки можно приблизить любую непрерывную функцию от одного аргумента $f(x)$ со сколь угодно большой точностью¹.

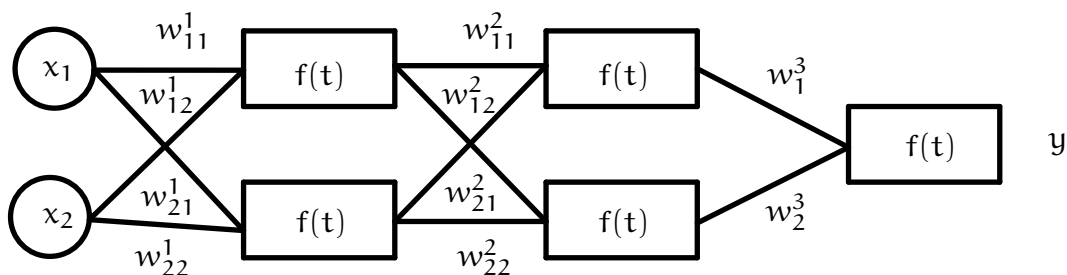
Hint: Вспомните, что любую непрерывную функцию можно приблизить с помощью кусочно-линейной функции (ступеньки). Осознайте как с помощью пары нейронов можно описать такую ступеньку. Соедините все ступеньки в сумму с помощью выходного нейрона.

Задача 7

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 .²

Задача 8

Дана нейросетка:



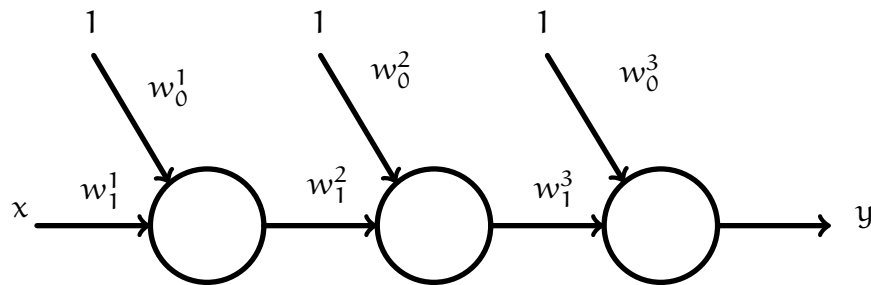
¹<http://neuralnetworksanddeeplearning.com/chap4.html>

²По мотивам книги Николенко "Глубокое обучение" (стр. 79)

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 .
4. Выглядит не очень оптимально, правда? Выпишите все производные в том виде, в котором их было бы удобно использовать для алгоритма обратного распространения ошибки, а затем, сформулируйте сам алгоритм.

Задача 9

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



В качестве функции активации используется сигмоид: $f(t) = \frac{e^t}{1+e^t}$. 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.

Hint: вам здорово поможет тот факт, что если $f(t) = \frac{e^t}{1+e^t}$, тогда $f'(t) = f(t)(1 - f(t))$. Это можно заодно и доказать.

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.

- а) Сколько параметров предстоит оценить Маше? Сколько наблюдений вы бы на её месте использовали?
- б) 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?
- в) Что делать Маше, если она хочет решать задачу классификации на K классов?