

# Тятя! Тятя! Нейросети заменили продавца!

Ппилиф Ульянкин

[https://github.com/FUlyankin/neural\\_nets\\_prob](https://github.com/FUlyankin/neural_nets_prob)

## Листочек 2: что выплёвывает нейросеть

Плюют в душу обычно те, кому не удалось в неё влезть.

*Пацанский наблик категории Б*

### Упражнение 1 (про сигмоиду)

Любую s-образную функцию называют сигмодой. Наиболее сильно прославилась под таким названием функция  $f(t) = \frac{e^t}{1+e^t}$ . Слава о ней добралась до Маши и теперь она хочет немного поисследовать её свойства<sup>1</sup>.

- а. Что происходит при  $t \rightarrow +\infty$ ? А при  $t \rightarrow -\infty$ ?
- б. Как связаны между собой  $f(t)$  и  $f(-t)$ ?
- в. Как связаны между собой  $f'(t)$  и  $f'(-t)$ ?
- г. Как связаны между собой  $f(t)$  и  $f'(t)$ ?
- д. Найдите  $f(0)$ ,  $f'(0)$  и  $\ln f(0)$ .
- е. Найдите обратную функцию  $f^{-1}(t)$ .
- ж. Как связаны между собой  $\frac{d \ln f(t)}{dt}$  и  $f(-t)$ ?
- з. Постройте графики функций  $f(t)$  и  $f'(t)$ .
- и. Говорят, что сигмоида — это гладкий аналог единичной ступеньки. Попробуйте построить на компьютере графики  $f(t)$ ,  $f(10 \cdot t)$ ,  $f(100 \cdot t)$ ,  $f(1000 \cdot t)$ . Как они себя ведут?

### Упражнение 2 (про logloss)

У Маши три наблюдения, первое наблюдение — кит, остальные — муравьи. Киты кодируются  $y_i = 1$ , муравьи —  $y_i = 0$ . В качестве регрессоров Маша берёт номера наблюдений  $x_i = i$ . После этого Маша оценивает логистическую регрессию с константой. В качестве функции потерь используются логистические потери.

---

<sup>1</sup>Часть задач украдена отсюда: [https://github.com/bdemeshev/mlearn\\_pro](https://github.com/bdemeshev/mlearn_pro)

- а. Выпишите для данной задачи функцию потерь, которую минимизирует Маша.
- б. При каких оценках коэффициентов логистической регрессии эта функция достигает своего минимума?

**Hint:** Изобразите наблюдения на числовой прямой и подумайте как должна будет вести себя идеальная сигмоида. Обратите внимание, что выборка линейно-разделима.

### Упражнение 3 (про softmax)

Маша чуть внимательнее присмотрелась к своему третьему наблюдению и поняла, что это не муравей, а бобёр. Теперь ей нужно решать задачу классификации на три класса. Она решила использовать для этого нейросеть с softmax-слоем на выходе.

Маша уже обучила нейронную сетку и хочет построить прогнозы для двух наблюдений. Слой, который находится перед softmax выдал для этих двух наблюдений следующий результат:  $(1, -2, 0)$  и  $(0.5, -1, 0)$ .

- а. Чему равны вероятности получить кита, муравья и бобра для этих двух наблюдений?
- б. Пусть первым был кит, а вторым бобёр. Чему будет равна logloss-ошибка?
- в. Пусть у Маши есть два класса. Она хочет выучить нейросеть. Она может учить нейронку с одним выходом и сигмоидой в качестве функции активации либо нейронку с двумя выходами и softmax в качестве функции активации. Как выходы этих двух нейронок взаимосвязаны между собой?
- г. Объясните, почему softmax считают сглаженным вариантом  $\arg \max$ .
- д. Докажите, что  $\text{softmax}(z + c) = \text{softmax}(z)$ , где  $c$  — какая-то константа, прибавленная ко всем выходам слоя. Как этот факт позволяет сделать softmax численно устойчивым?

### Упражнение 4 (про разные выходы)

Та, в чьих руках находится лёрнинг, решила немного поэкспериментировать с выходами из своей сетки.

- а. Маша решила, что хочет решать задачу классификации на два класса и получать на выходе вероятность принадлежности к первому. Что ей надо сделать с последним слоем сетки?
- б. Маша хочет решать задачу классификации на  $K$  классов. Что ей делать с последним слоем?
- в. Маша хочет спрогнозировать рейтинг фильма на "Кинопоиске". Он измеряется по шкале от 0 до 10 и принимает любое непрерывное значение. Как Маша может приспособить для этого свою нейронку?
- г. У Маши есть куча новостей. Каждая новость может быть спортивной, политической или экономической. Иногда новость может относиться сразу к нескольким категориям. Как Маше собрать нейросеть для решения этой задачи? Как будет выглядеть функция ошибки?

- д. У Маши есть картинки с уточками и чайками. Маша хочет научить нейросеть искать на картинке птицу, обводить её в прямоугольник (bounding box), а затем классифицировать то, что попало в прямоугольник. Как должен выглядеть выход из такой нейросети? Как должна выглядеть функция потерь?
- е. Маша задумалась, как можно спрогнозировать число людей в кафе так, чтобы на выходе сетка всегда прогнозировала целое число. Надо ли менять функцию потерь?