1. Опишите этапы построения линейных классификаторов. Чем они отличаются и чем схожи?

Линейный классификатор — алгоритм классификации, основанный на построении линейной разделяющей поверхности. В случае двух классов разделяющей поверхностью является гиперплоскость, которая делит пространство признаков на два полупространства. В случае большего числа классов разделяющая поверхность кусочно-линейна (как сумма выпуклых функций, которая тоже является выпуклой функцией).

Один из самых простых линейных классификаторов получается на основе регрессии вот таким образом:

$$a(\vec{x}) = sign(\vec{w}^T x),$$

где

- $\vec{x}$ – вектор признаков примера (вместе с единицей);
- $\vec{w}$ – веса в линейной модели (вместе со смещением w_0);
- $sign(\bullet)$ – функция "сигнум", возвращающая знак своего аргумента;
- $a(\vec{x})$ – ответ классификатора на примере $\vec{x}$.

Логистическая регрессия является частным случаем линейного классификатора, но она обладает хорошим "умением" – прогнозировать вероятность p_+ отнесения примера x_i к классу "+":

Посмотрим, как логистическая регрессия будет делать прогноз $p_+ = P\left(y_i = 1 \mid \vec{x_i}, \vec{w}\right)$ (пока считаем, что веса $\vec{w}$ мы как-то получили (т.е. обучили модель), далее разберемся, как именно).

- **Шаг 1.** Вычислить значение $w_0 + w_1 x_1 + w_2 x_2 + \ldots = \vec{w}^T \vec{x}$. (уравнение $\vec{w}^T \vec{x} = 0$ задает гиперплоскость, разделяющую примеры на 2 класса);

- **Шаг 2.** Вычислить логарифм отношения шансов: $\log(OR_+) = \vec{w}^T \vec{x}$.

- **Шаг 3.** Имея прогноз шансов на отнесение к классу "+" – OR_+, вычислить p_+ с помощью простой зависимости:

$$p_+ = \frac{OR_+}{1 + OR_+} = \frac{\exp^{\vec{w}^T \vec{x}}}{1 + \exp^{\vec{w}^T \vec{x}}} = \frac{1}{1 + \exp^{-\vec{w}^T \vec{x}}} = \sigma(\vec{w}^T \vec{x})$$

В правой части мы получили как раз сигмоид-функцию. Построения линейных классификаторов во многом различаются, но в каждом из них неизменным остается принцип разделяющей поверхности.