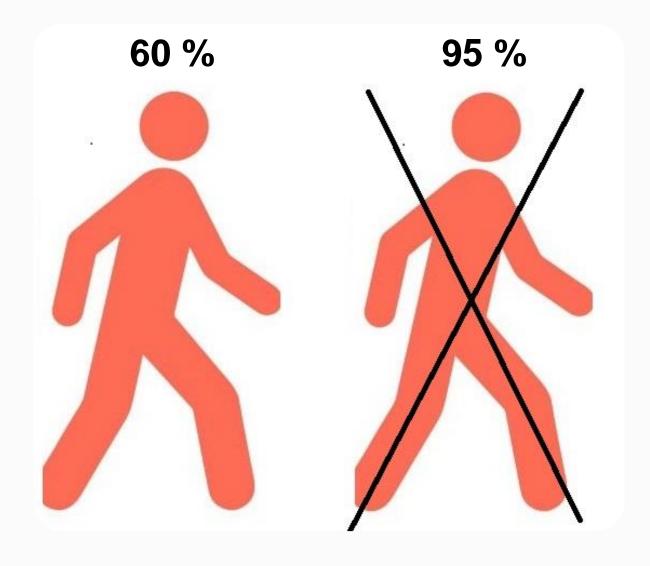
Логистическая регрессия

Елена Кантонистова

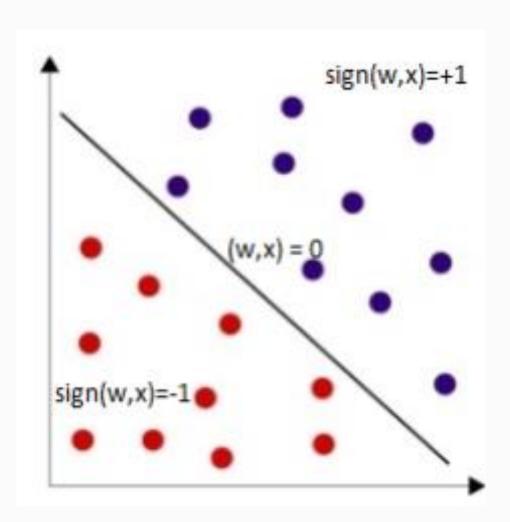
Skillbox

Мягкая классификация

Нам хочется знать не только ответ, но и вероятность ответа.

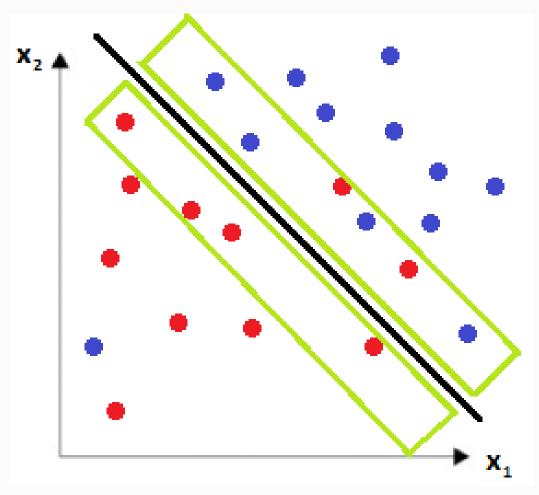


Бинарный линейный классификатор



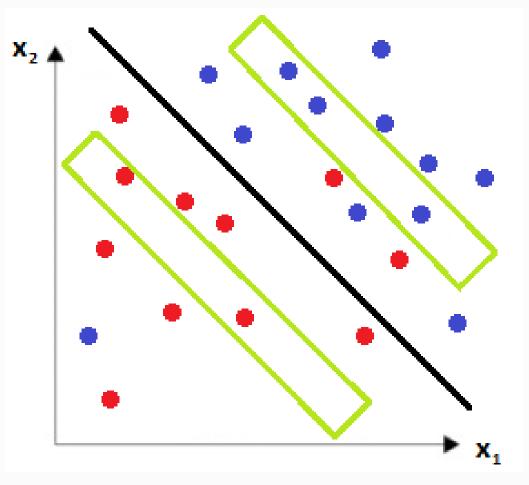
Изображение: <u>Два множества, разделённых</u> линейно (ru.wikipedia.org)

Уверенность классификатора



Изображение: работа спикера

Уверенность классификатора



Изображение: работа спикера

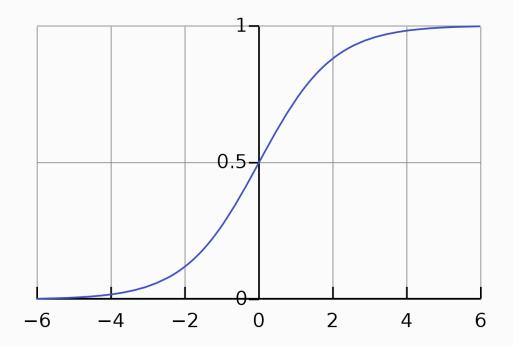
Чего хотим

Хотим, чтобы бинарный классификатор умел выдавать не только класс, но и **вероятность того**, **что объект относится к положительному классу**.

Сигмоида

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

Сигмоида переводит произвольное действительное число (расстояние от объекта до разделяющей границы) в отрезок [0; 1].

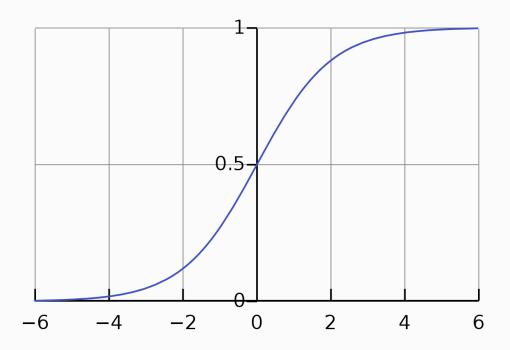


Изображение: <u>Логистическая функция активации</u> (ru.wikipedia.org)

Предсказываем вероятность

$$a(x) = \sigma(w, x) = \frac{1}{1 + e^{-(w, x)}}$$

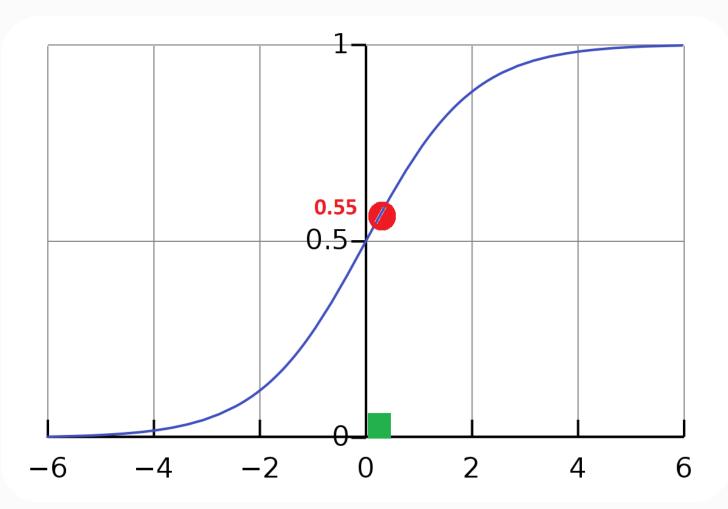
Применение сигмоиды к скалярному произведению позволяет получить число из отрезка [0; 1].



Изображение: <u>Логистическая функция активации</u> (ru.wikipedia.org)

Как это работает?

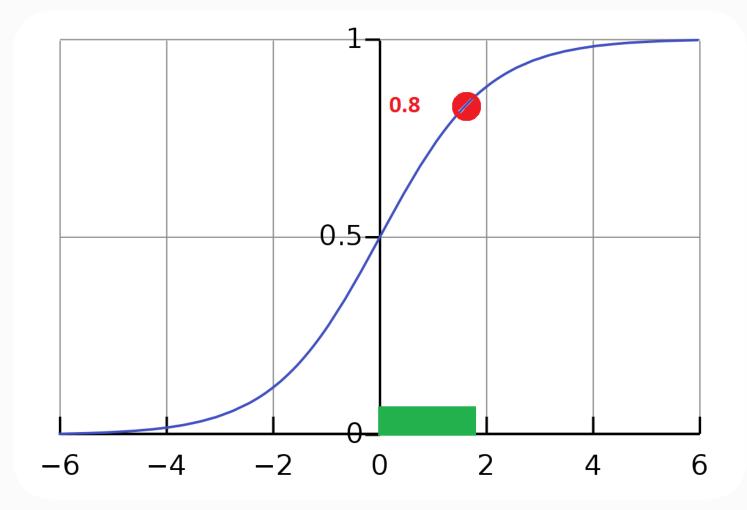
Объект близко к разделяющей прямой и (w, x) > 0



Изображение: Логистическая функция активации (ru.wikipedia.org)

Как это работает?

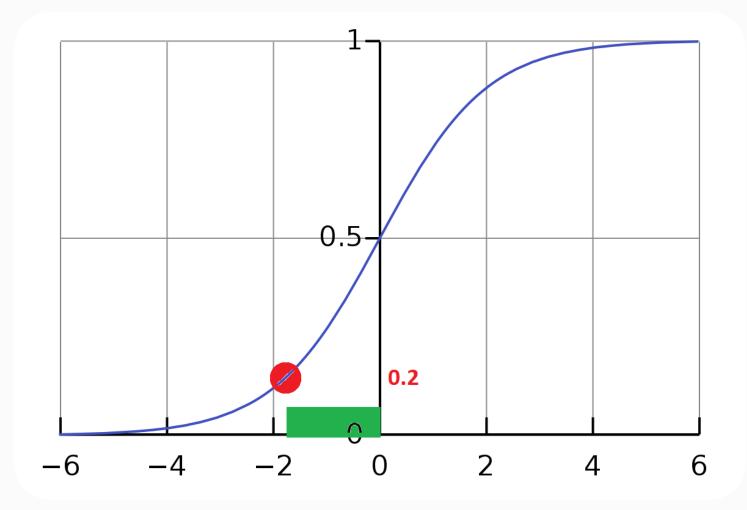
Объект в положительном классе и далеко от разделяющей прямой.



Изображение: Логистическая функция активации (ru.wikipedia.org)

Как это работает?

Объект в отрицательном классе.



Изображение: Логистическая функция активации (ru.wikipedia.org)

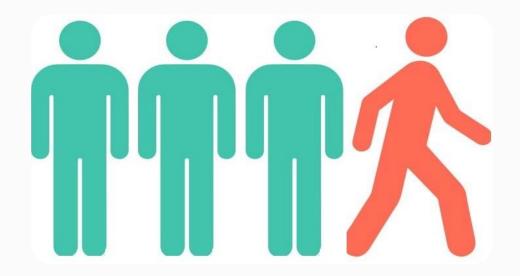
Логистическая регрессия

$$a(x) = \frac{1}{1 + e^{-(w,x)}}$$

Логистическая регрессия для модели оттока

- Получаем не только класс (уйдёт клиент или нет),
 но и вероятность того, что клиент уйдёт из компании
- Получаем хорошо интерпретируемую модель, например:

$$a(x) = \sigma(1 + \mathbf{10} \cdot \mathbf{возраст} - \mathbf{5} \cdot \mathbf{задолженность})$$



Перевод вероятности в класс

$$a(x) = \frac{1}{1 + e^{-(w,x)}}$$

- Если a(x) ≥ 0.5, то класс +1
- Если a(x) < 0.5, то класс −1

Итог

Логистическая регрессия — это линейный классификатор, который умеет предсказывать вероятности классов.

$$a(x) = \frac{1}{1 + e^{-(w,x)}}$$