

Линейный классификатор и нейрон

Ксемидов Борис Сергеевич

Chillers

16 апреля, 2020

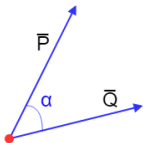
Скалярное произведение

Скалярное произведение выражается следующими образами:

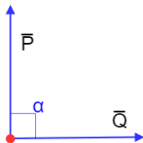
$$(a, b) = a_1 b_1 + a_2 b_2 + \dots + a_n b_n$$

$$(a, b) = |a||b|\cos\alpha$$

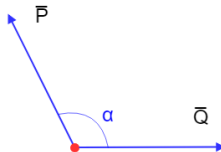
Скалярное произведение



$\alpha < 90^\circ$, острый угол
 $\vec{P} \cdot \vec{Q} > 0$



$\alpha = 90^\circ$, прямой угол
 $\vec{P} \cdot \vec{Q} = 0$



$\alpha > 90^\circ$, тупой угол
 $\vec{P} \cdot \vec{Q} < 0$

Рисунок: Вариации значений скалярного произведения

sgn

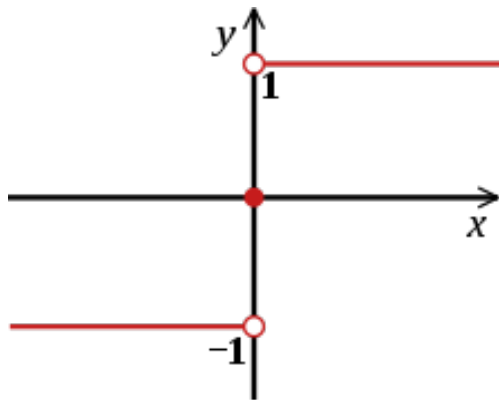


Рисунок: $\text{sgn} x$

Линейный классификатор

Есть два класса: $Y = -1, 1$.

Тогда можно построить линейный классификатор вида:

$$a(x, w) = \operatorname{sgn}\left(\sum_{i=1}^m w_i x_i - b\right)$$

Обучение линейного классификатора

Обучение линейного классификатора сводится к минимизации следующего функционала:

$$L(w) = \sum_{i=1}^m [a(x_i, w) \neq y_i]$$

или

$$L(w) = \sum_{i=1}^m [a(x_i, w)y_i < 0]$$

Аппроксимация минимизируемого функционала

Вместо ранее указанного функционала часто для оптимизации берут непрерывные функции:

- $\log_2(1 + e^{-M})$ - логистическая аппроксимация
- e^{-M} - экспоненциальная аппроксимация

Аппроксимация минимизируемого функционала

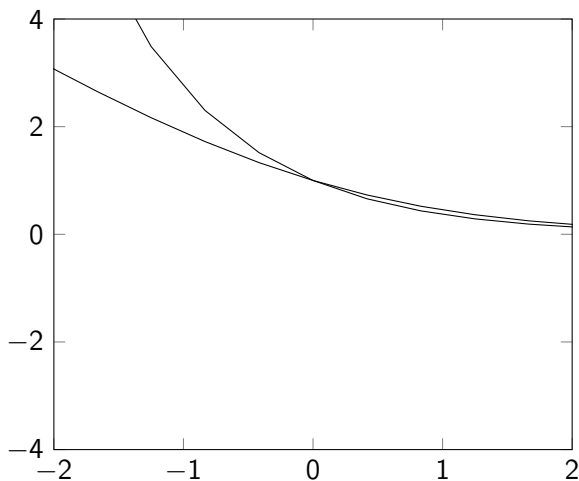


Рисунок: Графики минимизируемого функционала

Пример линейного классификатора

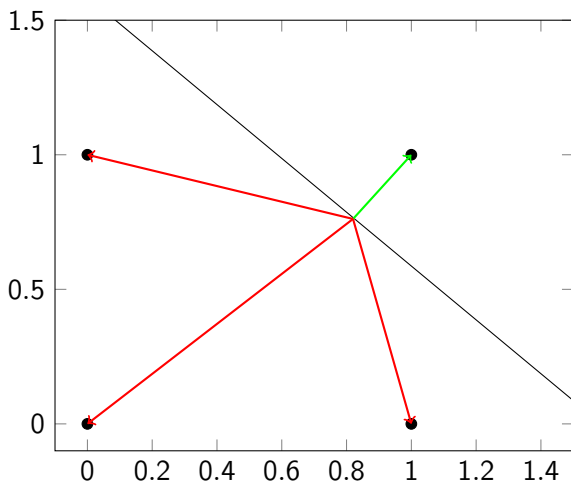


Рисунок: Логическая операция "И"

Нейрон

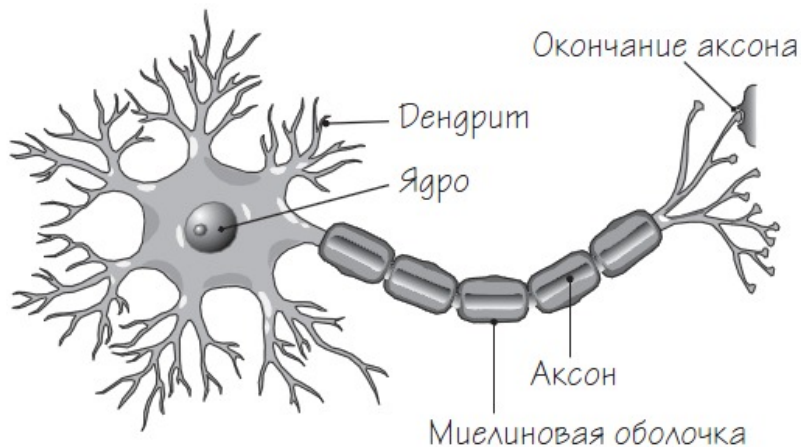


Рисунок: Строение нейрона

Модель нейрона МакКаллока-Питтса

Модель нейрона МакКаллока-Питтса выглядит следующим образом.

$$a(x, w) = \sigma\left(\sum_{i=1}^m x_i\right),$$

Функция активации

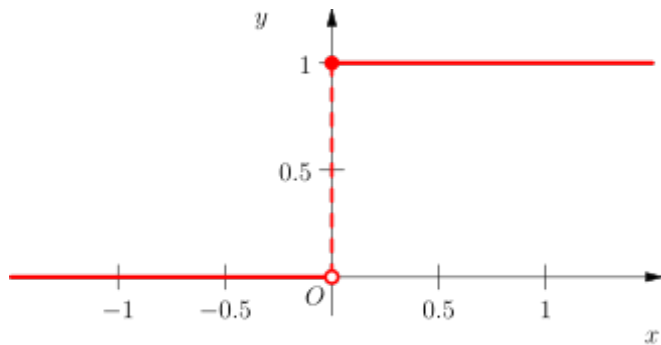


Рисунок: Функция Хевисайда

Модель нейрона МакКаллока-Питтса

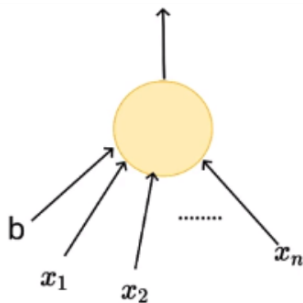


Рисунок: Нейрон МакКаллока-Питтса

Персептрон Розенблатта

Персептрон представляет собой линейный классификатор:

$$a(x, w) = \sigma\left(\sum_{i=1}^m w_i x_i - b\right)$$

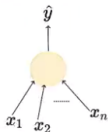
Обучение персептрона Розенблатта

Персептрон обучается посредством правила Хэбба:

- если сигнал персептрона неверен и равен нулю, то необходимо увеличить веса тех входов, на которые была подана единица.
- если сигнал персептрона неверен и равен единице, то необходимо уменьшить веса тех входов, на которые была подана единица.

Отличия нейрона МП от персептрона

MP Neuron

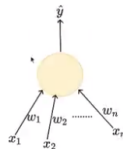


$$\hat{y} = 1 \text{ if } \sum_{i=1}^n x_i \geq b$$

$$\hat{y} = 0 \text{ otherwise}$$

- ☹ Boolean inputs
- ☹ Linear
- ☹ Inputs are not weighted
- 😊 Adjustable threshold

Perceptron



$$\hat{y} = 1 \text{ if } \sum_{i=1}^n w_i x_i \geq b$$

$$\hat{y} = 0 \text{ otherwise}$$

- 😊 Real inputs
- ☹ Linear
- 😊 Weights for each input
- 😊 Adjustable threshold

Рисунок: Преимущества и недостатки персептрона и нейрона МП

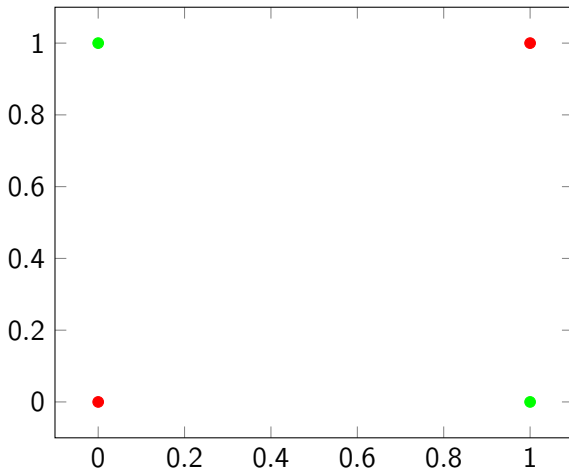
XOR

a	b	c
0	0	0
0	1	1
1	0	1
1	1	0

Таблица: Таблица истинности XOR

Персептрон и XOR

Как будет выглядеть разделяющая прямая в этом случае?



Другие функции активации

- $\sigma(x) = \frac{1}{1+e^{-x}}$ - сигмоида
- $\sigma(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ - тангенсальная
- $\sigma(x) = \ln(1 + e^x)$ - SoftPlus

Спасибо за внимание!