

Лекция1

...

Expressml

Приложения нейронных сетей

- Компьютерное зрение
- Распознавание речи
- Анализ текста
- Управление сложными системами
- Анализ разнородных данных
- Сжатое представление данных
- Artistic view
- Автоматические автомобили

Одна из первых задач MNIST

~ 32x32

Распознавание рукописных цифр



CIFAR

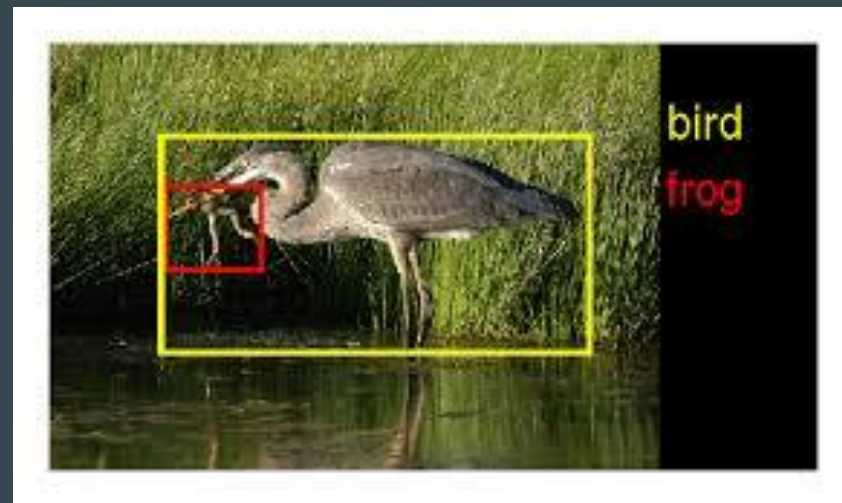
 32×32

Распознавание объектов на маленьких картинках



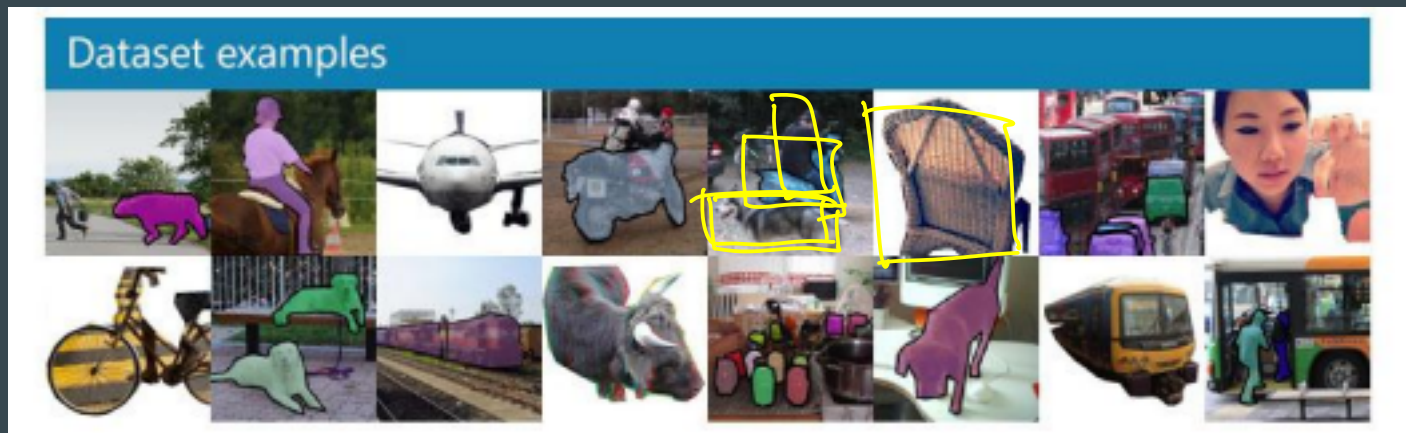
ImageNet — 1000 1.5 млн
— 21k ~15 млн

Распознавание объектов на картинках



COCO ~ 120к картинок

- Детектирование объектов
- Локализация объектов
- Сегментация объектов
- Сопоставление предложения картинке
- ...



Синтаксический анализ текста



Мнемоническое правило

Нейросеть хорошо справляется с тем, с чем
хорошо справляется человек

План на сегодня

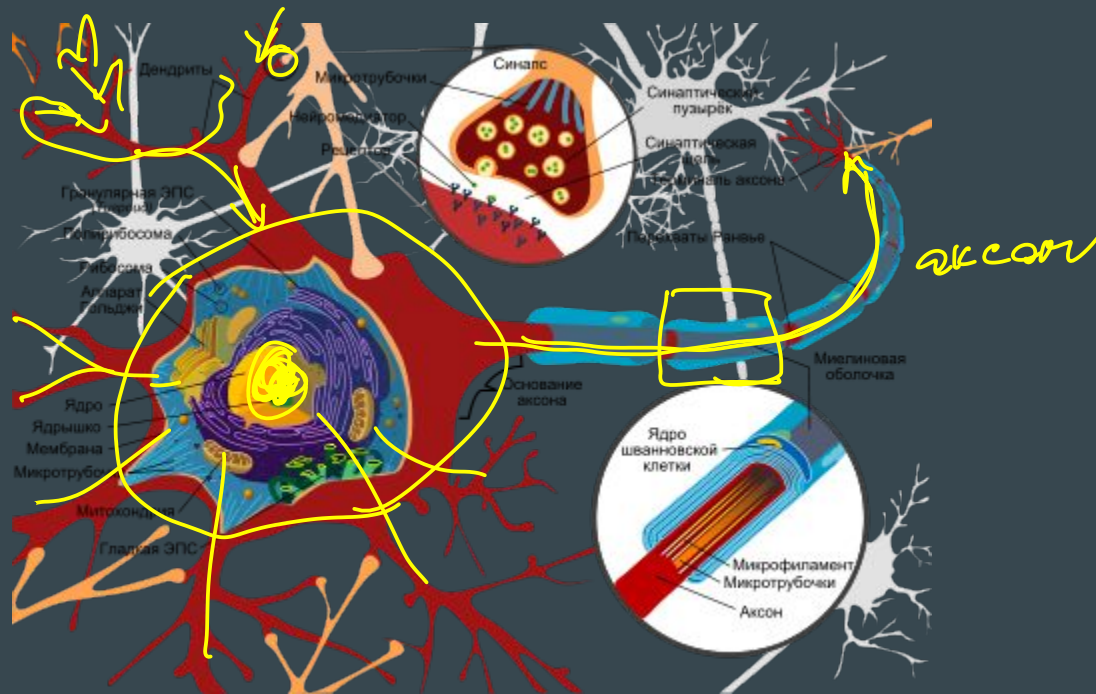
- Математическая модель нейрона
 - Биологический нейрон
 - Математическая модель
 - Функции активации
- Нейросети, выполняющие функции логических элементов
 - NOT
 - AND
 - OR
 - XOR
- Двуслойная нейронная сеть, почему она нелинейная

План на сегодня

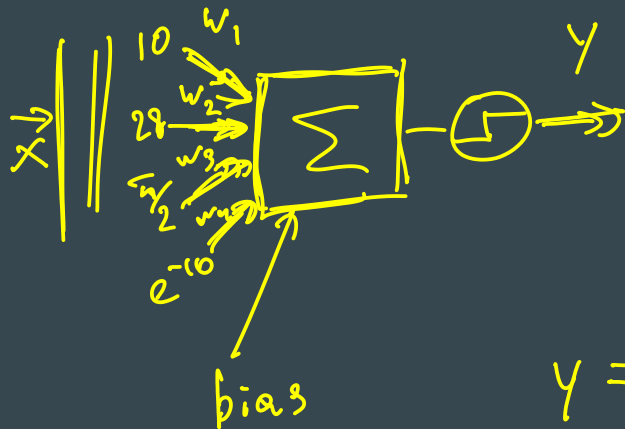
- Что нужно для тренировки:
 - Архитектура
 - Функция потерь и метрики
 - Оптимизатор
- Типы задач, которые решаются при помощи нейронных сетей
 - Классификация
 - Регрессия
 - Ранжирование
 - Сжатие размерности

Математическая модель нейрона

Биологический нейрон



Математическая модель



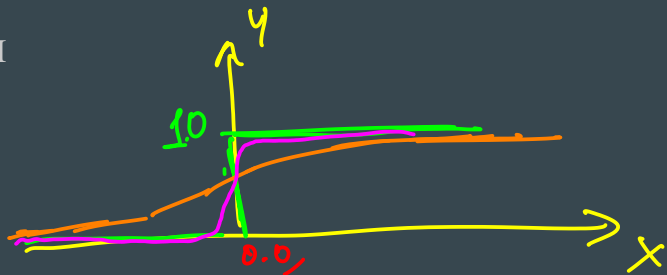
$$\text{thresh}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

$$\boxed{100.054}$$

$$y = \text{act} \left(\sum_{i=1}^N w_i \cdot x_i + \text{bias} \right)$$

Функции активации

- Пороговая



- Сигмоида

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

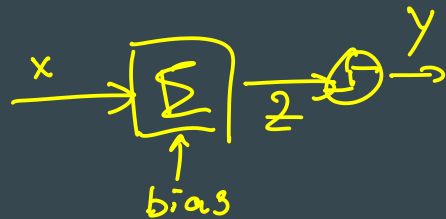
$$\sigma(x) \quad x \rightarrow -\infty = 0$$

$$x \rightarrow +\infty = 1$$

$$\sigma'(x) = \left(\frac{1}{1 + e^{-x}} \right)' = \frac{+e^{-x}}{(1 + e^{-x})^2}$$

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

Нейрон NOT



$$y = \text{act} \left(w \cdot x + b \right)$$

Th2 $w = -1$
 $b = 0.5$

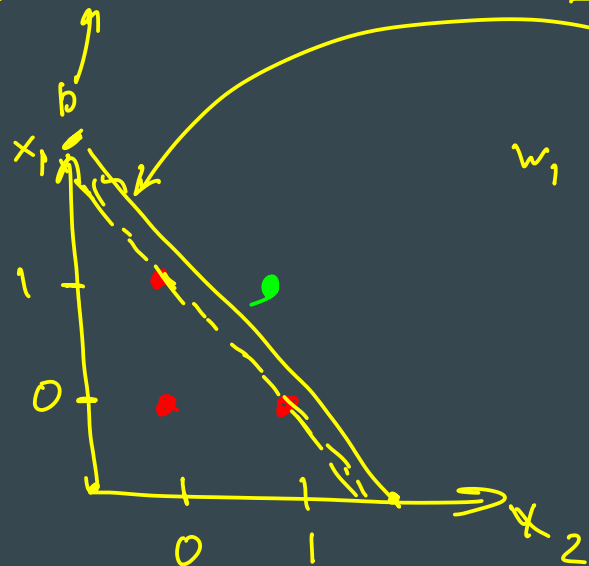
x	\bar{x}
1	0
0	1

Нейрон AND



$$y = \text{actf}(\underbrace{w_1 x_1 + w_2 x_2 + b}_{\text{линейная ф-ция}})$$

линейная
ф-ция



$$x_1 + x_2 = 1.5$$

$$\begin{matrix} x_1 & + & x_2 & - & 1.5 \\ \uparrow & & \uparrow & & \downarrow \\ -1 & & 2 & & b \end{matrix}$$

$w_1 = 1 \quad w_2 = 1$

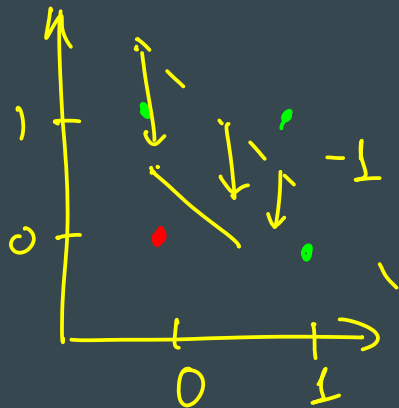
$$y = \text{thresh}(x_1 + x_2 - 1.5)$$

x_1	x_2	$x_1 \& x_2$
1	1	1
1	0	0
0	1	0
0	0	0

Нейрон OR

$$y = \text{act}(w_1 x_1 + w_2 x_2 + b)$$

x_1	x_2	
1	1	1
1	0	1
0	1	1
0	0	0

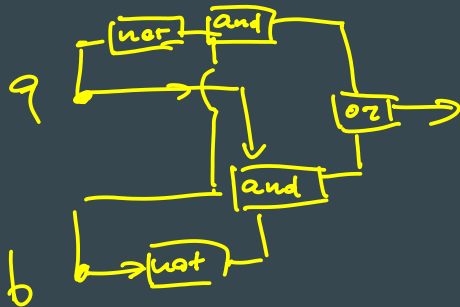
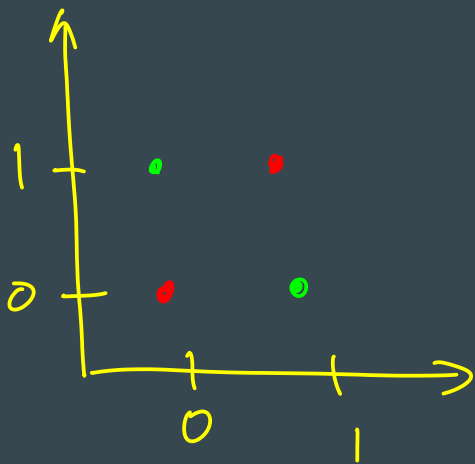


$$y = \text{thresh}(x_1 + x_2 - 0.5)$$

Нейрон XOR

$$a \text{ xor } b = (a \text{ and } \bar{b}) \text{ or } (\bar{a} \text{ and } b)$$

x_1	x_2	
1	1	0
0	1	1
1	0	1
0	0	0

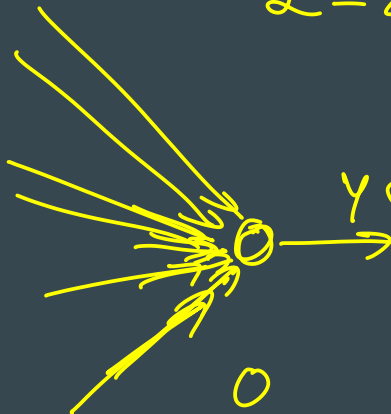


Нейронная сеть

Однослойная нейронная сеть, линейность



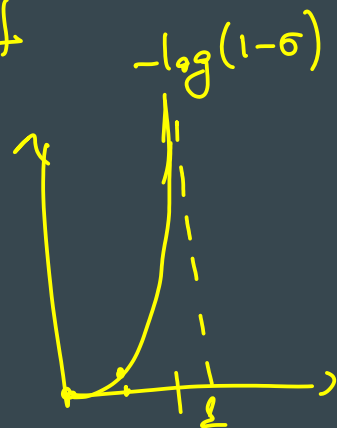
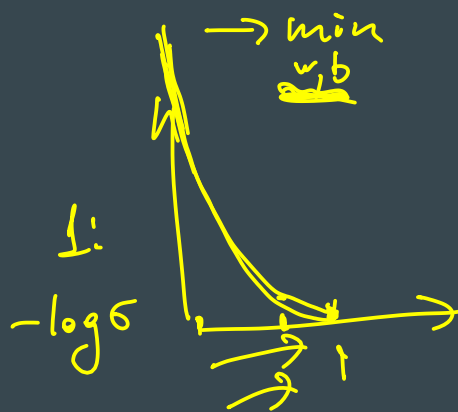
\vec{x}



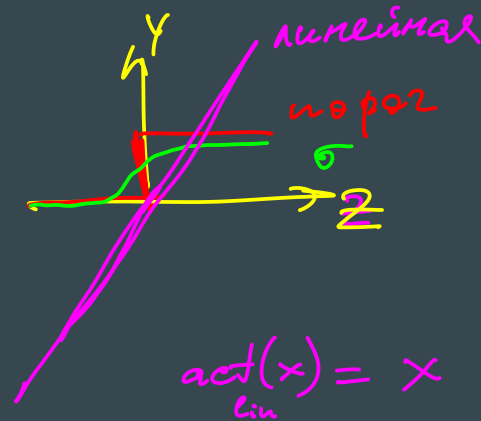
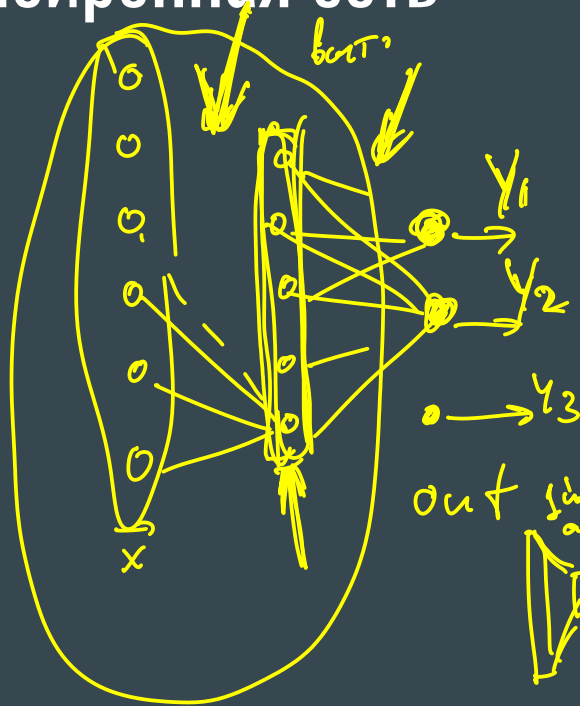
$$L = \sum_i BCE(y_i, \hat{x}_i) =$$

$$= \sum_i -y_i \log \sigma(z_i) - (1-y_i) \log(1-\sigma(z_i))$$

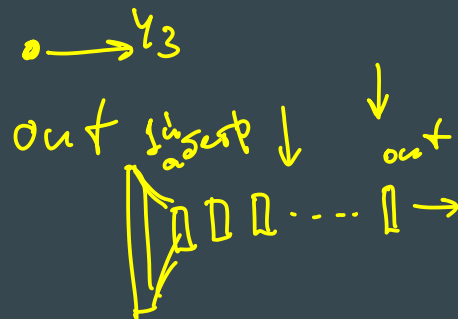
$$z = \sum_i w_i x_i + b$$



Двуслойная нейронная сеть



Th.



Двуслойная нейронная сеть

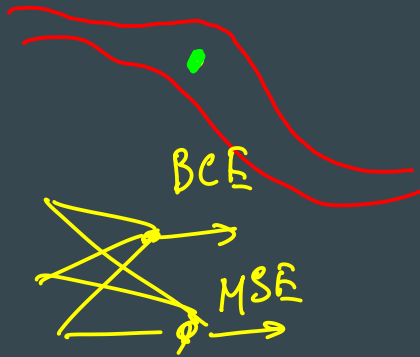
Больше слоёв...

Все что нужно :)

Архитектура



Функция потерь



$$L = L_1 + L_2$$

$$MSE = \frac{1}{N} \sum_{i=1}^N (\text{out}_i - y_i)^2$$

нейронный $\sigma(z)$

$$BCE = -y \log p - (1-y) \log (1-p)$$

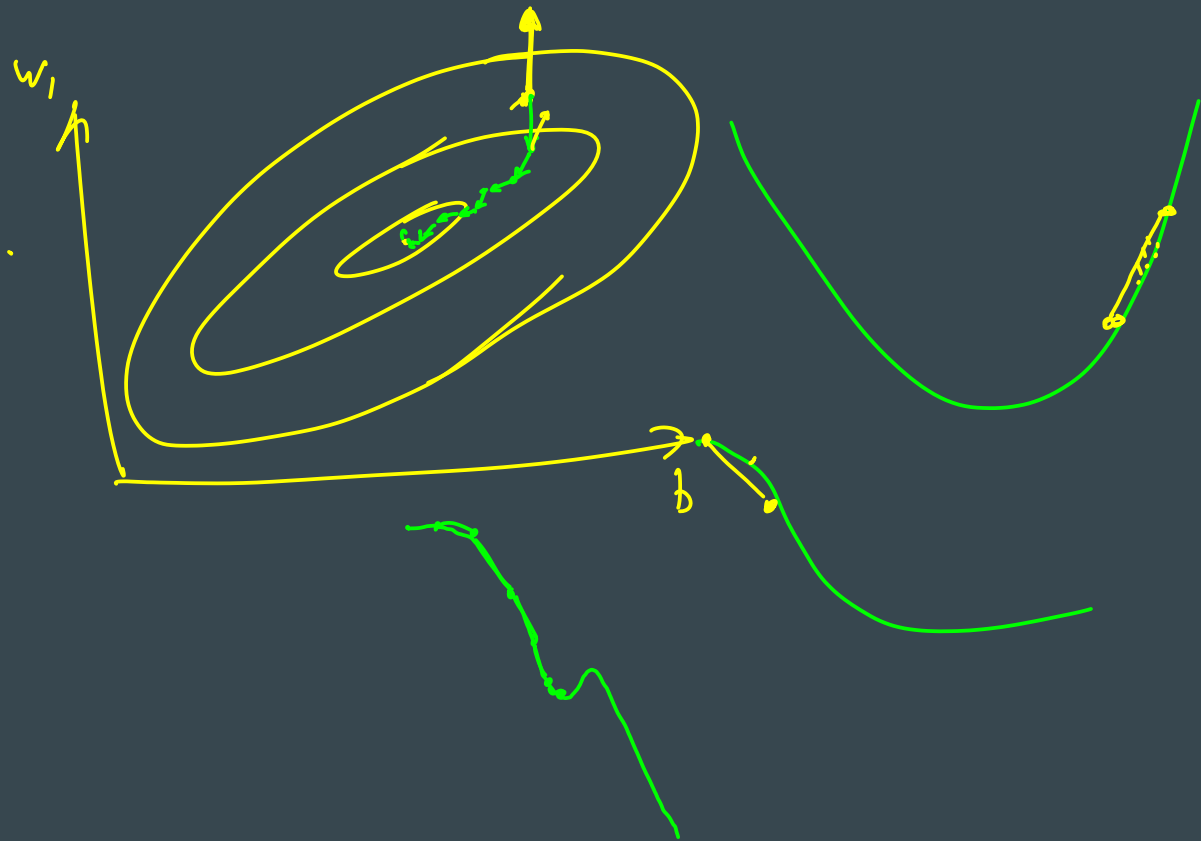
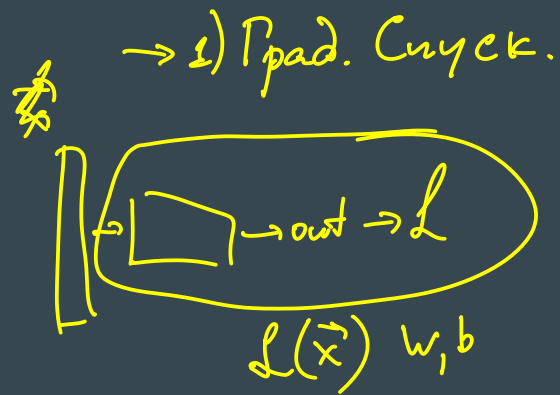
$$\frac{\partial BCE}{\partial z} = -\frac{y}{p} \cdot p' - \frac{(1-y)}{1-p} \cdot (-p') =$$

$$= -\frac{y}{\sigma} (1-\sigma) + \frac{(1-y)}{1-\sigma} \sigma =$$

$$= -y + y\sigma + \sigma - y\sigma = \sigma - y$$

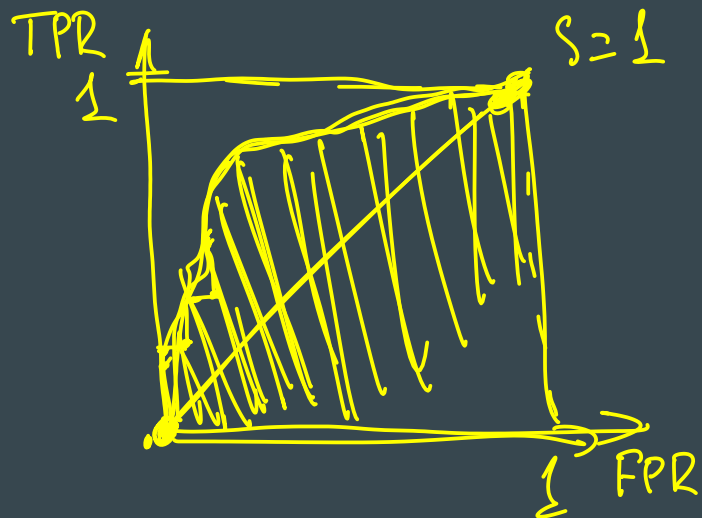


Оптимизатор

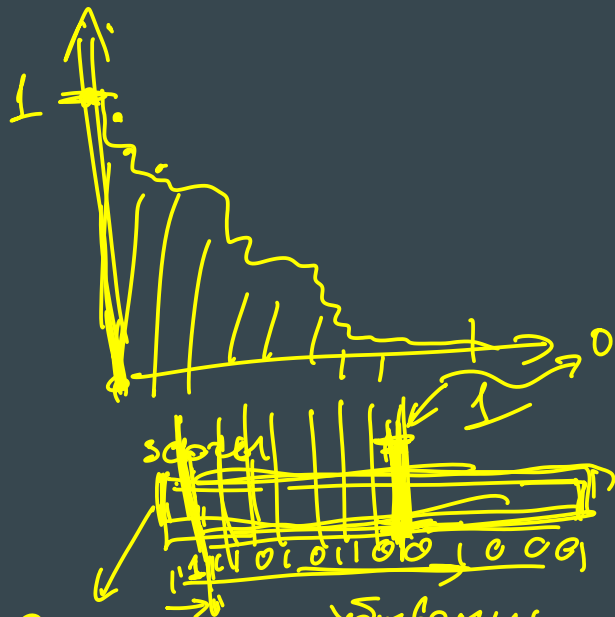


Метрики

ROC-AUC



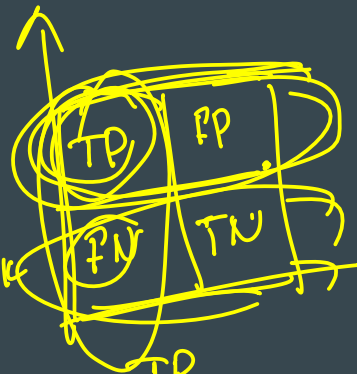
0-1



Prec.
Rec.

f-мера

PRC



$$P = \frac{TP}{TP + FP}$$

$$R = \frac{TP}{TP + FN}$$

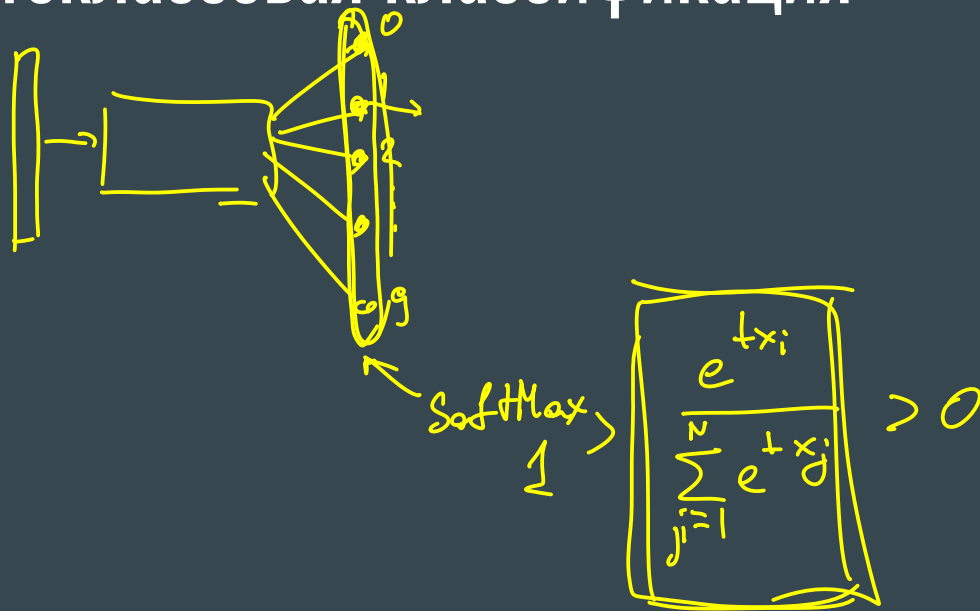
Типы задач, решаемые при помощи Нейронных Сетей

Бинарная классификация

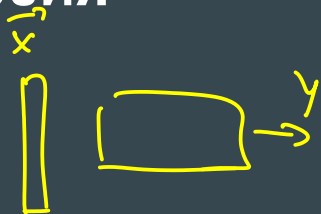


Многоклассовая классификация

$\boxed{1}$ $0 \dots 9$



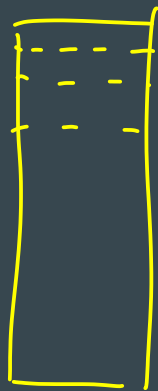
Регрессия



$$MSE \quad \sum (f(x) - y)^2$$

$$L1 \quad \sum |f(x) - y| \quad MAE$$

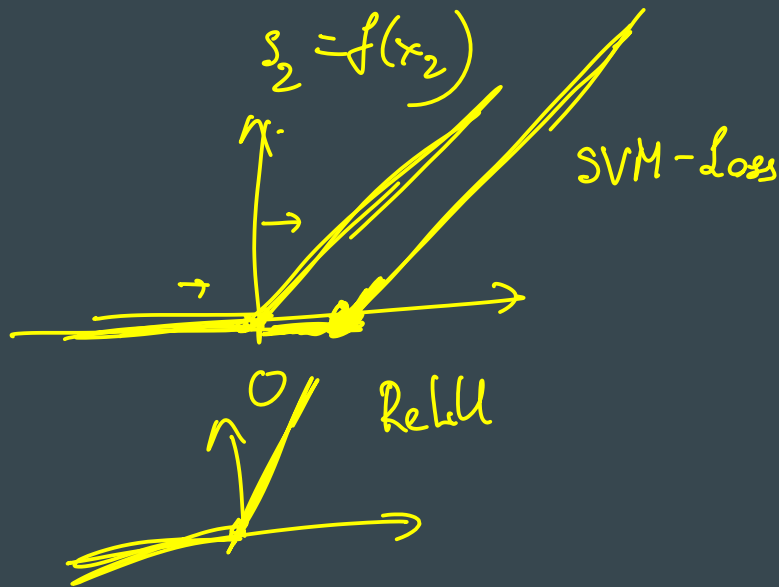
Ранжирование



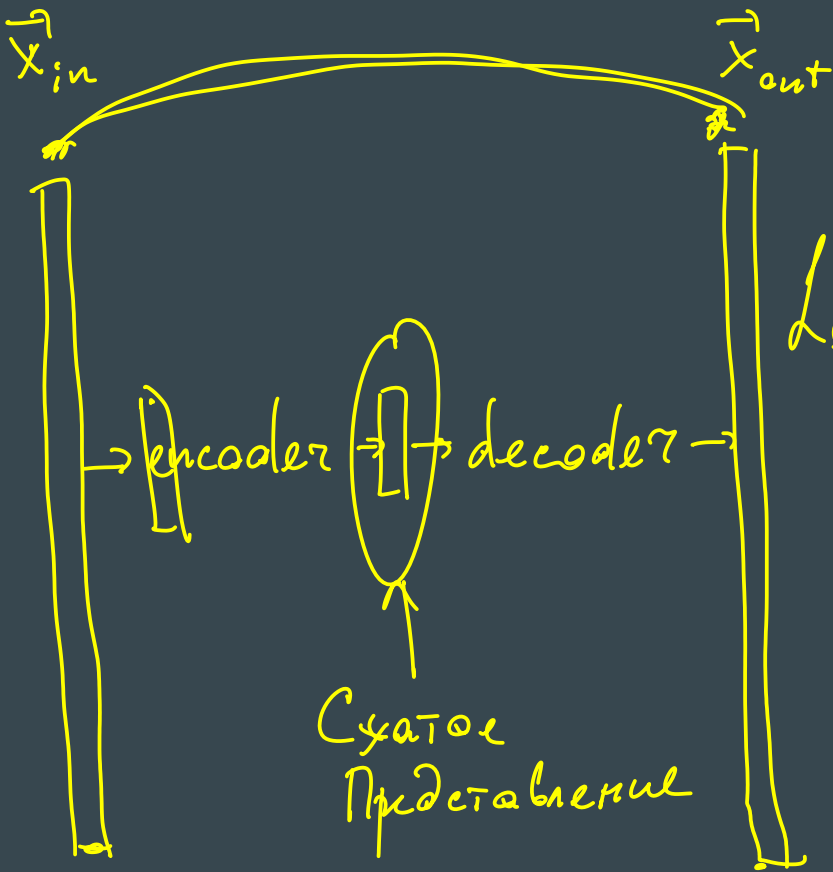
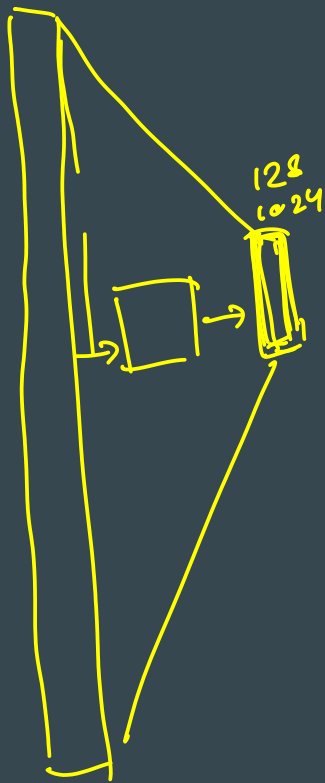
Margin Loss: $s_1 > s_2$
 $\max\{0, \frac{1}{m} - (s_1 - s_2)\}$

$s_1 = f(x_1) \quad m > 0$

$s_2 = f(x_2)$



Сжатие размерности



$$Loss = \sum (\vec{x}_{in} - \vec{x}_{out})^2$$

نہ

،

N независимых p -орн $(1-p)$ -ремка

n была орн

$$P = C_N^n \cdot p^n \cdot (1-p)^{(N-n)} = \text{Likelihood}(p) \rightarrow \max_p$$

$$\log \text{Likelihood} = \boxed{\log C_N^n} + n \cdot \log p + (N-n) \log(1-p) \rightarrow \max_p$$

$$\begin{aligned} & n \cdot \log p + (N-n) \log(1-p) \rightarrow \log \text{loss} \\ & -\frac{n}{N} \log p - \left(1 - \frac{n}{N}\right) \log(1-p) \rightarrow \min_p \end{aligned}$$

$x \in$ Гауссовский процесс

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

x_1, \dots, x_N

$$P_{\{x_1, \dots, x_N\}} = p(x_1)p(x_2) \cdot \dots \cdot p(x_N) = \left(\frac{1}{\sqrt{2\pi}}\right)^N \cdot \frac{1}{\sigma^N} \cdot \exp\left[-\sum_{i=1}^N \frac{1}{2\sigma^2} (x_i - \mu)^2\right]$$

$$\text{Log likelihood} = -N \log \sqrt{2\pi} - N \log \sigma - \sum_{i=1}^N \frac{1}{2\sigma^2} (x_i - \mu)^2 \rightarrow \max_{\mu, \sigma}$$