



ТЕХНОСФЕРА

Лекция 1 Основы нейронных сетей

Полыковский Даниил

12 февраля 2018 г.

Биологический нейрон

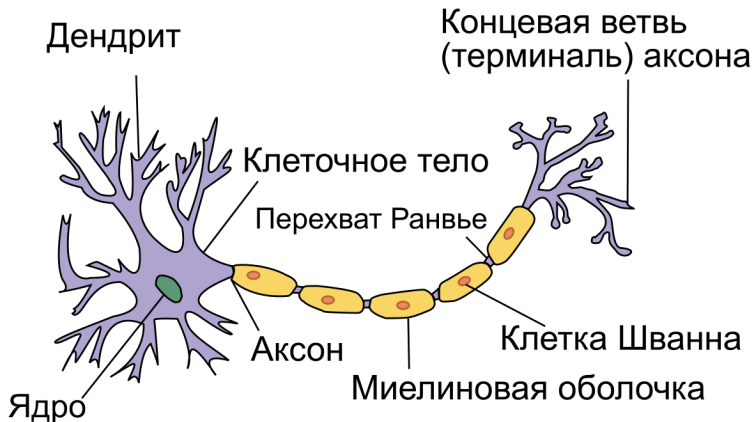


Рис.: Структура нейрона

Биологический нейрон

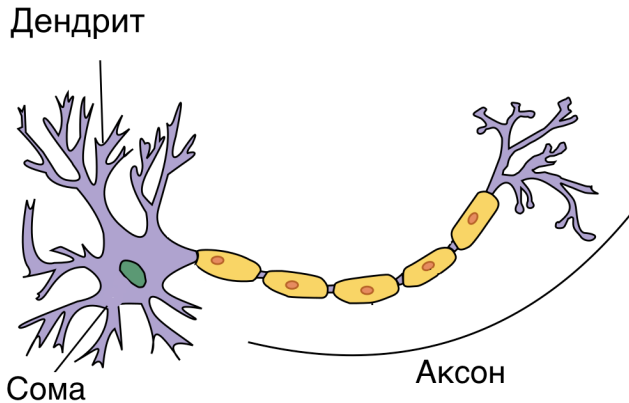


Рис.: Структура нейрона

Биологический нейрон

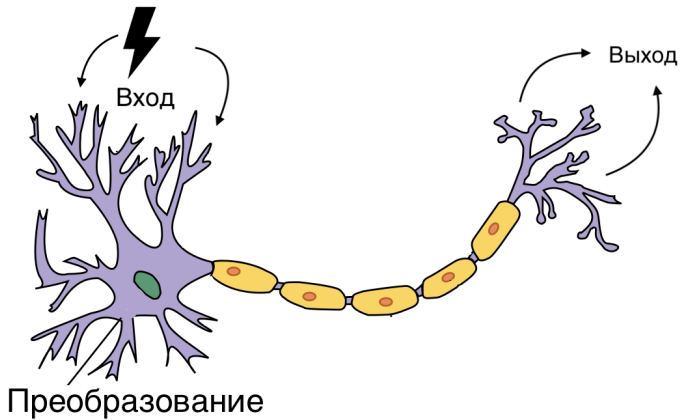


Рис.: Структура нейрона

Биологический нейрон

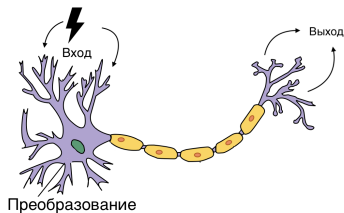


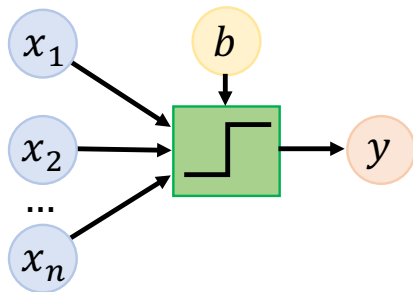
Рис.: Структура нейрона

Выходной сигнал посылается при достижении определенного уровня входного сигнала.

Модель:

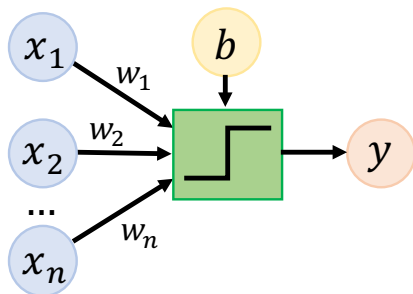
$$y = \begin{cases} 1, & \sum_{i=1}^N x_i > b \\ 0, & \text{иначе} \end{cases} = \mathbb{I}\left[\sum_{i=1}^N x_i > b\right]$$

Модель 1: Схема



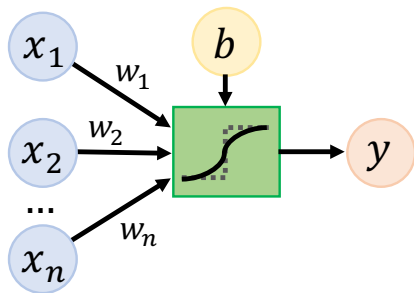
$$y = \mathbb{I}\left[\sum_{i=1}^N x_i > b\right]$$

Модель 2: Чувствительность нейронов



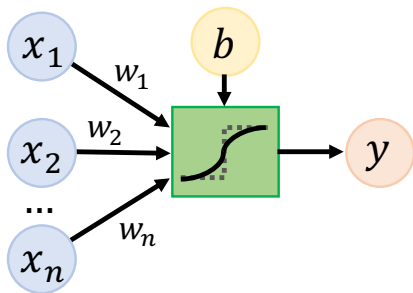
$$y = \mathbb{I}\left[\sum_{i=1}^N w_i x_i > b\right] = \mathbb{I}[w^T x > b]$$

Модель 3: Непрерывная активация



$$y = \sigma\left[\sum_{i=1}^N w_i x_i - b\right] = \sigma[w^T x - b]$$

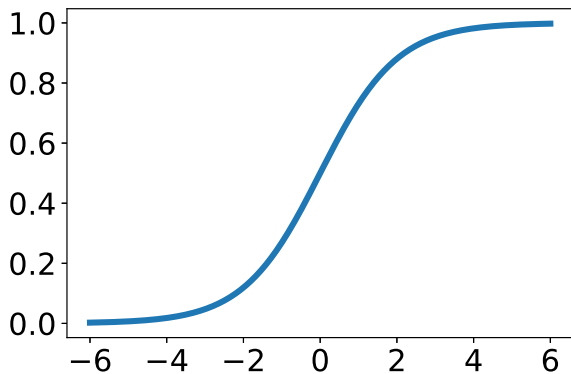
Модель 3: Непрерывная активация



$$y = \sigma\left[\sum_{i=1}^N w_i x_i - b\right] = \sigma[w^T x - b]$$

Параметры нейрона: w — веса, b — смещение.

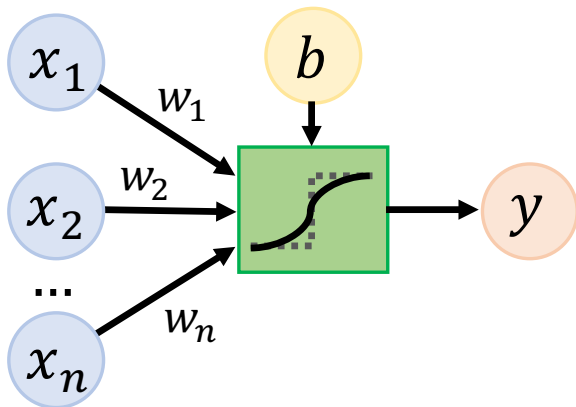
Функция активации



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

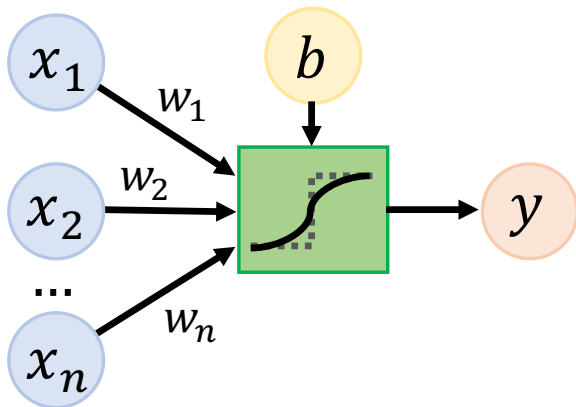
Полносвязные сети

Перцептрон



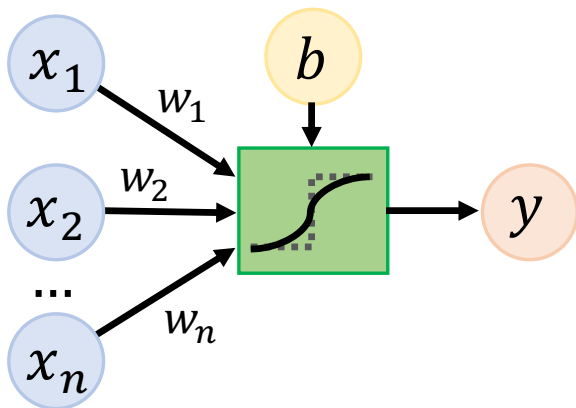
► Можем моделировать:

Перцептрон



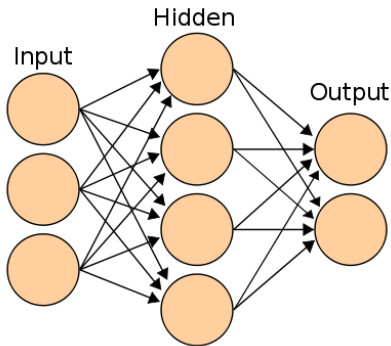
- ▶ Можем моделировать: NOT, AND, OR
- ▶ Не можем моделировать:

Перцептрон



- ▶ Можем моделировать: NOT, AND, OR
- ▶ Не можем моделировать: XOR

Сети с одним скрытым слоем



Теорема (универсальный аппроксиматор)¹

Любую непрерывную на компакте функцию можно равномерно приблизить нейронной сетью с одним скрытым слоем.

¹Отличная визуализация:

<http://neuralnetworksanddeeplearning.com/chap4.html>

Как обучить нейронную сеть?

Обучить нейронную сеть — подобрать значения всех настраиваемых параметров (веса и смещения).

Два этапа:

1. Задать функцию потерь \mathcal{L}
2. Подобрать веса, минимизирующие \mathcal{L}

Регрессия

Хотим предсказать непрерывную величину y для объекта X (задача регрессии).

Имеется выборка $(X_1, t_1), \dots, (X_n, t_n)$. Пусть сеть предсказала y на объекте с правильной меткой t . Функция потерь: $\delta = t - y$

- ▶ Mean Squared Error (MSE): δ^2
- ▶ Mean Absolute Error (MAE): $|\delta|$
- ▶ Разная стоимость пере- и недопрогноза:
 $\delta^2(a \cdot \mathcal{I}[\delta < 0] + b \cdot \mathcal{I}[\delta \geq 0])$

Классификация: Negative log-likelihood

Хотим предсказать класс объекта. Функция потерь:

$$- \sum_{c=1}^C \mathbb{I}[t = c] \log p_c$$

p_1, \dots, p_C — предсказания вероятностей от нейросети: $\sum_{c=1}^C p_c = 1$

- ▶ Как обеспечить условие $\sum_{c=1}^C p_c = 1$?

Классификация: Negative log-likelihood

Хотим предсказать класс объекта. Функция потерь:

$$- \sum_{c=1}^C \mathbb{I}[t = c] \log p_c$$

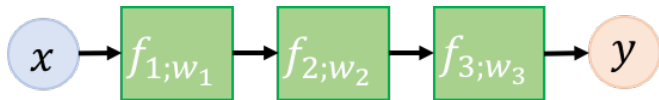
p_1, \dots, p_C — предсказания вероятностей от нейросети: $\sum_{c=1}^C p_c = 1$

- ▶ Как обеспечить условие $\sum_{c=1}^C p_c = 1$?
- ▶ $p_i = \frac{e^{y_i}}{\sum_{k=1}^C e^{y_k}} \leftarrow \text{Softmax}$ («мягкий» максимум)

Отдельный случай: бинарная классификация.

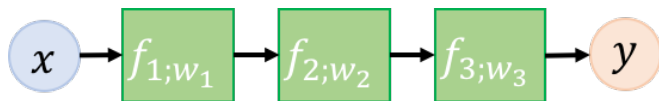
- ▶ $\mathcal{L} = -t \log(p) - (1 - t) \log(1 - p)$
- ▶ Достаточно одного выхода нейросети, пропущенного через σ

Обучение



$$y = f_3(f_2(f_1(x; w_1); w_2); w_3)$$

Обучение

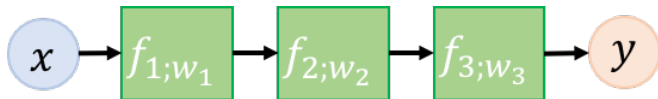


$$y = f_3(f_2(f_1(x; w_1); w_2); w_3)$$

$$y = \sigma \left(W_3 [\sigma (W_2 [\sigma (W_1 x + b_1)]) + b_2] + b_3 \right)$$

$$\mathcal{L} = ||y - t||^2 \rightarrow \min_{W_1, W_2, W_3, b_1, b_2, b_3}$$

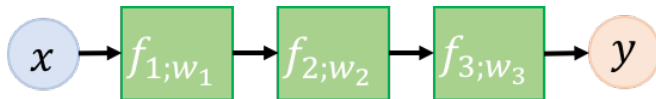
Обучение



Градиентный спуск:

1. $W^0 \leftarrow$ Начальное приближение : $W_i \sim \mathcal{N}(0, 0.1)$, $b_i = 0$
2. WHILE not converged:
3. $W^k \leftarrow W^{k-1} - \eta \nabla_W \mathcal{L}$

Обучение

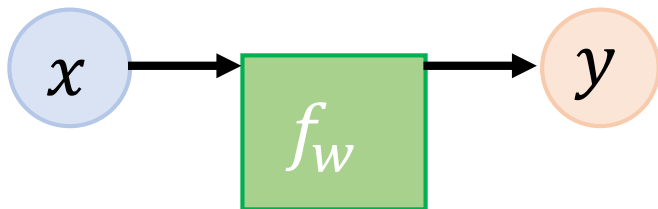


Градиентный спуск:

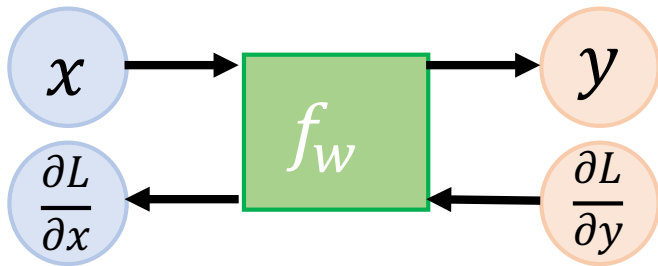
1. $W^0 \leftarrow$ Начальное приближение : $W_i \sim \mathcal{N}(0, 0.1)$, $b_i = 0$
2. WHILE not converged:
3. $W^k \leftarrow W^{k-1} - \eta \nabla_W \mathcal{L}$

$y = \sigma \left(W_3 [\sigma (W_2 [\sigma (W_1 x + b_1)]) + b_2] + b_3 \right)$: Сложность вычисления градиента растет с увеличением размера сети

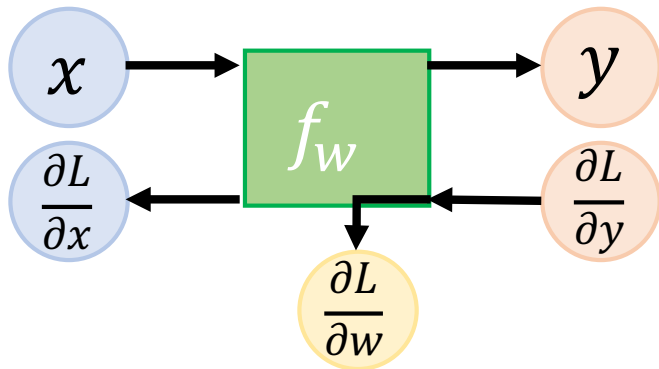
Back propagation



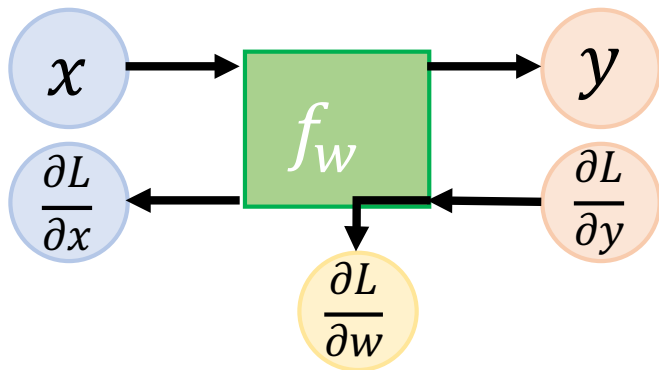
Back propagation



Back propagation

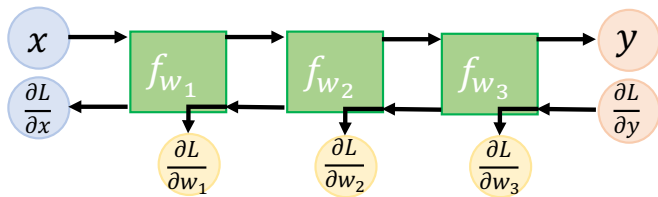


Back propagation

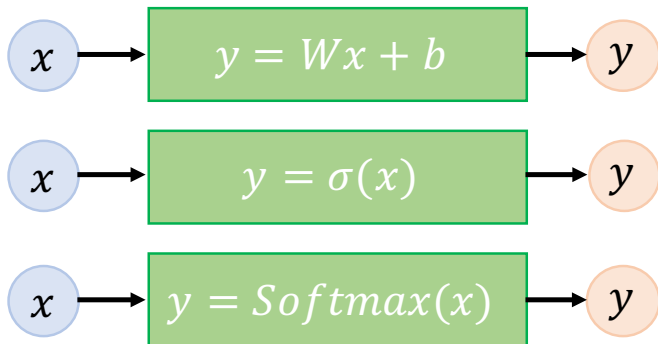


$$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial y} \frac{\partial y}{\partial w}, \quad \frac{\partial L}{\partial x} = \frac{\partial L}{\partial y} \frac{\partial y}{\partial x}$$

Back propagation



Building blocks



Gradient checking

Можно проверить корректность реализации, сравнив:

- ▶ Посчитанный градиент
- ▶ Численный градиент:

$$\frac{\partial L}{\partial w} \approx \frac{L(w + \epsilon) - L(w - \epsilon)}{2\epsilon}$$

Применение нейронных сетей

Neural networks vs Human writing

80322-4129 80206
40004 14310
37879 05453
5502 75216
35460 44209

Рис.: Zip codes

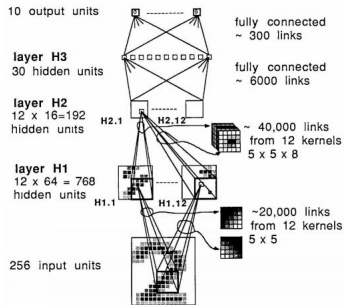


Рис.: Network



The South Korean professional Go player Lee Sedol reviews the match after finishing against Google's artificial-intelligence program, AlphaGo.

Lee Jin-man / AP

How Google's AlphaGo Beat a Go World Champion

Imagenet

Numbers in brackets: (the number of synsets in the subtree).

- ImageNet 2011 Fall Release (32326)
 - plant, flora, plant life (4486)
 - phytoplankton (2)
 - microflora (0)
 - crop (9)
 - cash crop (0)
 - catch crop (0)
 - cover crop (0)
 - field crop (4)
 - field corn (3)
 - dent corn, Zea mays ir
 - flint corn, flint maize,
 - soft corn, flour corn, s
 - root crop (0)
 - endemic (0)
 - holophyte (0)
 - non-flowering plant (0)
 - plantlet (0)
 - wilding (141)
 - ornamental (1)
 - pot plant (0)
 - acrogen (0)
 - apomict (0)
 - aquatic (0)
 - cryptogam (1)
 - annual (0)
 - biennial (0)
 - perennial (1)
 - escape (0)
 - hygrophyte (0)
 - neophyte (0)

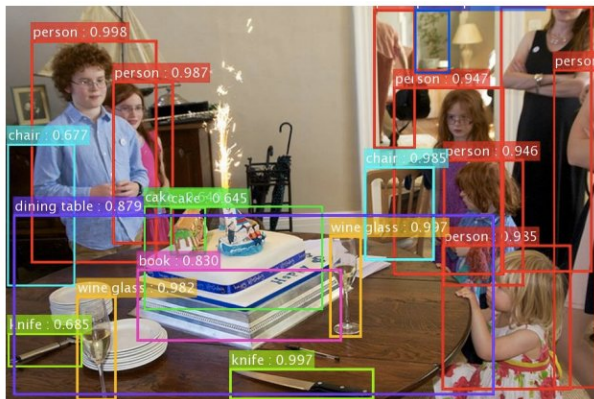
Still working...

[Treemap Visualization](#)
[Images of the Synset](#)
[Downloads](#)

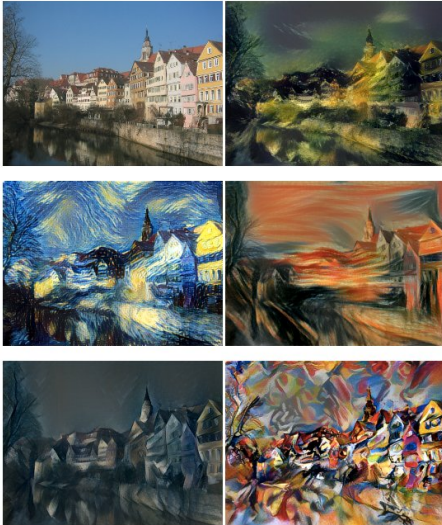
*Images of children synsets are not included. All images shown are thumbnails. Images may be subject to copyright.

[Prev](#)
[1](#)
[2](#)
[3](#)
[4](#)
[5](#)
[6](#)
[7](#)
[8](#)
[9](#)
[10](#)
[11](#)
[12](#)
[Next](#)

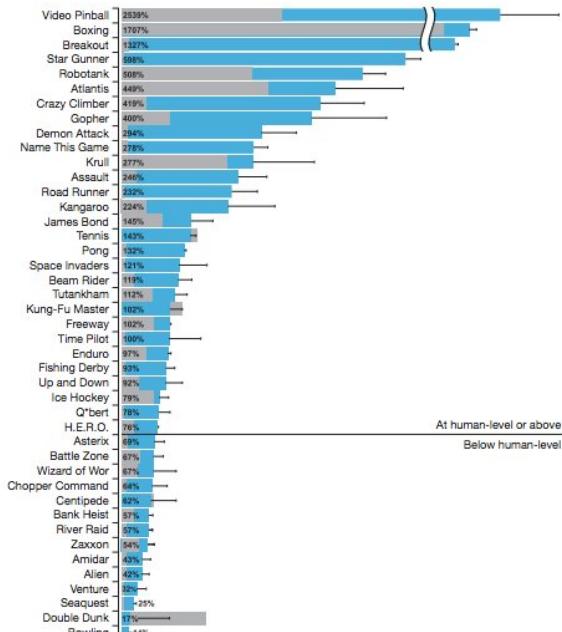
Classification: object detection



Artistic Style



Neural networks vs Games



DCGAN



Рис.: Faces

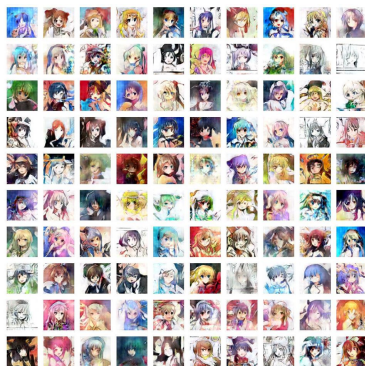
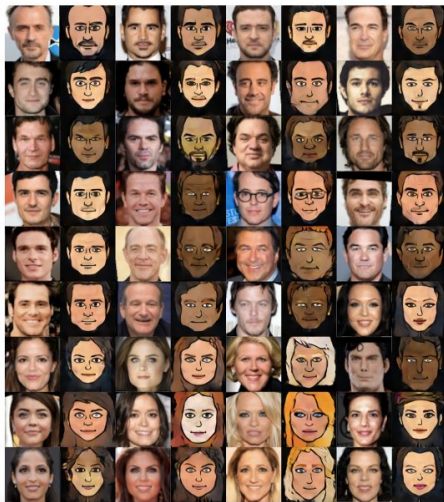


Рис.: Anime

Более хорошие результаты:

<https://www.youtube.com/watch?v=X0xxPcy5Gr4>

Cross-domain



Организационные вопросы

Структура курса

- ▶ Часть 1: Общая теория
 - ▶ 7 лекций (1.5 часа)
 - ▶ 7 семинаров (1.5 часа) [21 балл]
 - ▶ Коллоквиум [15 баллов]
- ▶ Часть 2: Применение
 - ▶ 7 лекций (1.5 часа)
 - ▶ 7 семинаров (1.5 часа) [21 балл]
 - ▶ Коллоквиум [15 баллов]
- ▶ Исследовательский проект [30 баллов]

Структура курса

- ▶ Часть 1: Общая теория
 - ▶ 7 лекций (1.5 часа)
 - ▶ 7 семинаров (1.5 часа) [21 балл]
 - ▶ Коллоквиум [15 баллов]
- ▶ Часть 2: Применение
 - ▶ 7 лекций (1.5 часа)
 - ▶ 7 семинаров (1.5 часа) [21 балл]
 - ▶ Коллоквиум [15 баллов]
- ▶ Исследовательский проект [30 баллов]

Разбалловка:

- ▶ [0, 50) — «неудовлетворительно»
- ▶ [50, 70) — «удовлетворительно»
- ▶ [70, 80) — «хорошо»
- ▶ 80+ — «отлично»

Структура курса

- ▶ Общение: Slack, #_dm2_neural_networks
- ▶ Лекторы:
 - ▶ Даниил Поликовский (@danpol, daniil.polykovskiy@gmail.com)
 - ▶ Кузьма Храбров (@kuzma, k.khrabrov@corp.mail.ru)
- ▶ Задание посылается на почту `networks.sphere@gmail.com`
- ▶ Тема письма: "[NN18, Задание X], Фамилия Имя".
- ▶ Reading group: по понедельникам, 21:00-22:00, ауд. 582 (<http://bit.ly/readml>)

Семинар

№1: Матричные производные

- ▶ $\frac{\partial}{\partial x} x^T A x = ?$
- ▶ $\frac{\partial}{\partial A} x^T A x = ?$
- ▶ $\frac{\partial}{\partial x} \|Ax + b\|^2 = ?$
- ▶ $\frac{\partial}{\partial A} \text{tr}(AB) = ?$

№2: Линейная регрессия

- ▶ Выпишите функционал для линейной регрессии
- ▶ Найдите оптимальное значение весов

№3: Дифференцирование NLL

- ▶ Выпишите значение связки Softmax + NLL
- ▶ Выведите градиент $\frac{\partial}{\partial y} \text{NLL}(\text{Softmax}(y); t)$

№4: Логистическая регрессия

- ▶ $p(y|x) = \sigma(y[w^T x])$
- ▶ Выпишите функцию правдоподобия
- ▶ Выпишите градиент логарифма функции правдоподобия
- ▶ Как ведет себя градиент на правильно классифицированных объектах?
- ▶ Как ведет себя градиент на неправильно классифицированных объектах?

Вопросы

