



DEEP LEARNING

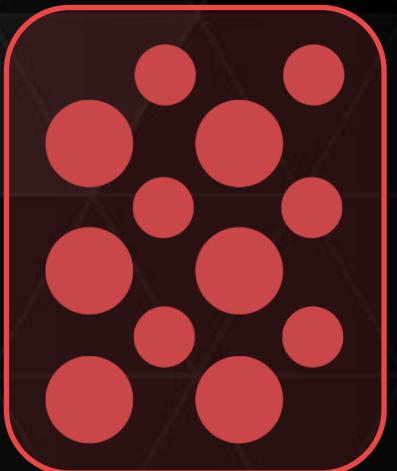
КАК РАБОТАЮТ И ОБУЧАЮТСЯ СОВРЕМЕННЫЕ НЕЙРОННЫЕ СЕТИ

Дмитрий Коробченко

МАШИННОЕ ОБУЧЕНИЕ

МАШИННОЕ ОБУЧЕНИЕ

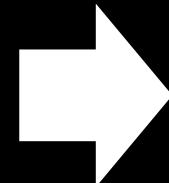
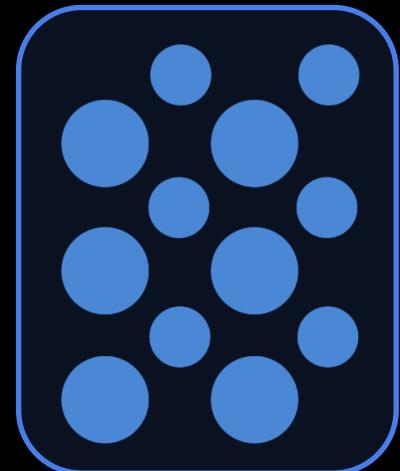
Вход



Модель

Параметры

Выход

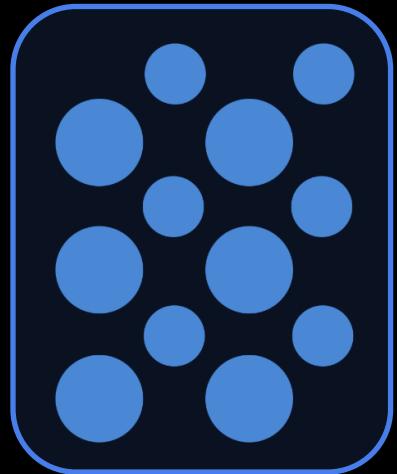
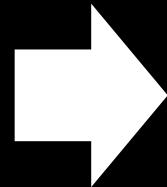
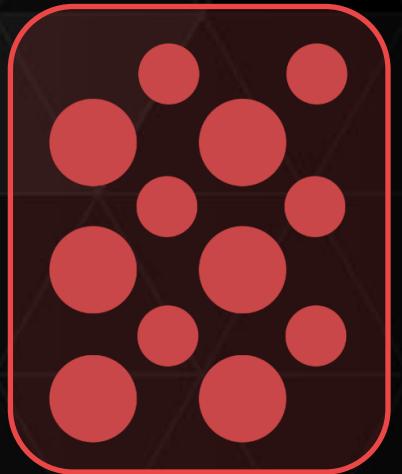


ОБУЧЕНИЕ С УЧИТЕЛЕМ



ВЫВОД (INFERENCE)

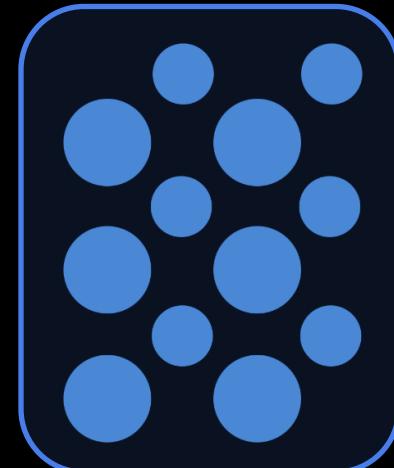
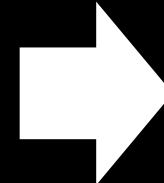
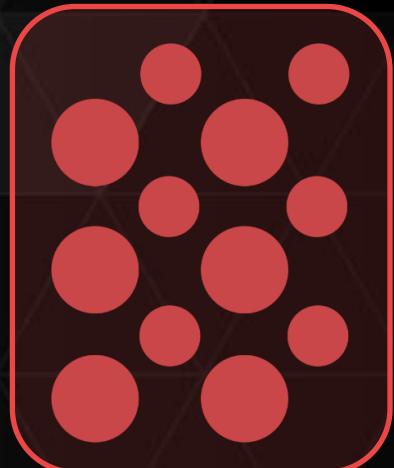
Новые данные



Ответ

ВЫВОД (INFERENCE)

Новые данные



Модель

Параметры

Обобщение

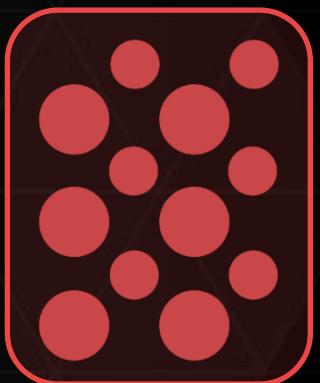
Накопленный опыт

Закономерности в данных

Запоминание обучающей выборки

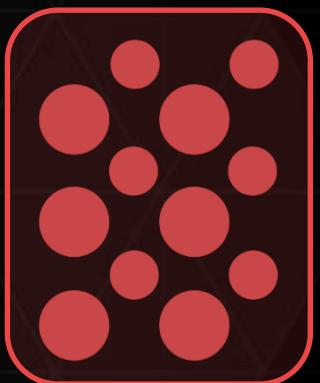
ПРИЗНАКИ

Признаковое
описание


$$\{x_1, x_2, x_3 \dots\}$$

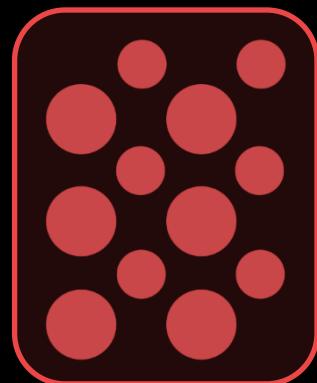
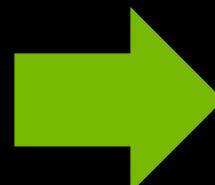
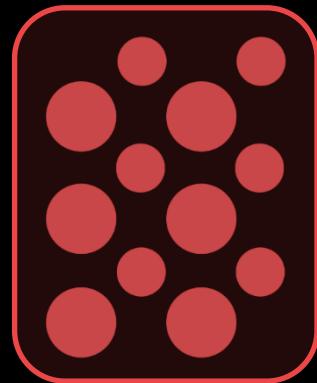
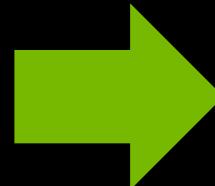
ПРИЗНАКИ

Признаковое
описание

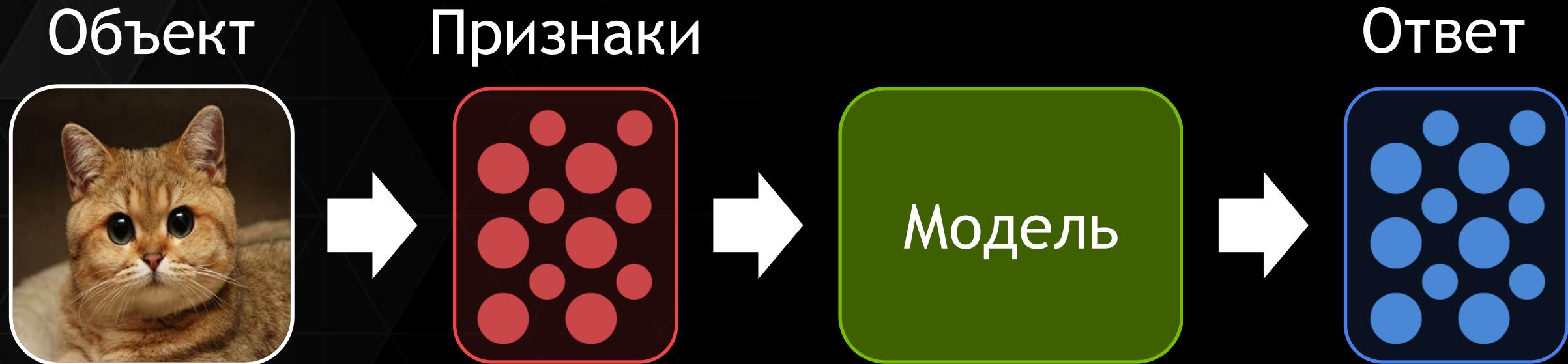


$$\{x_1, x_2, x_3 \dots\}$$

Извлечение признаков

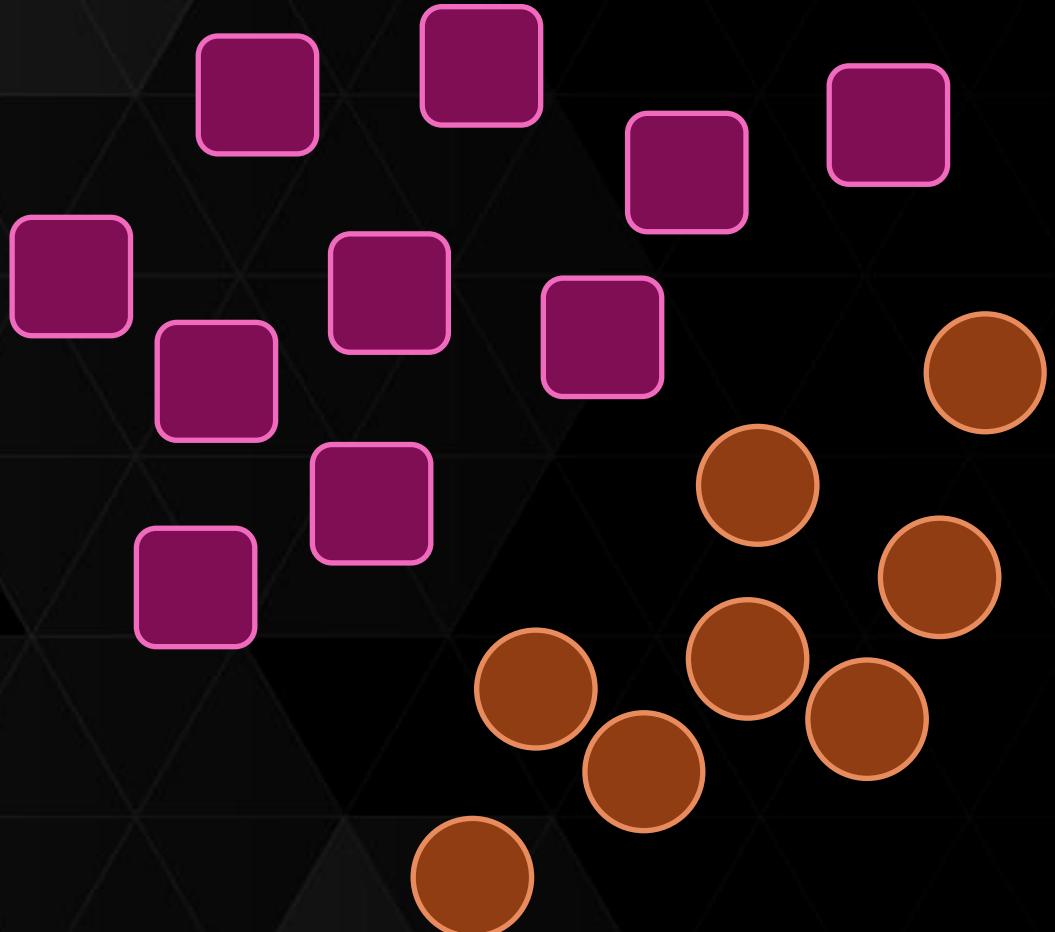


ИЗВЛЕЧЕНИЕ ПРИЗНАКОВ + ИНФЕРЕНС

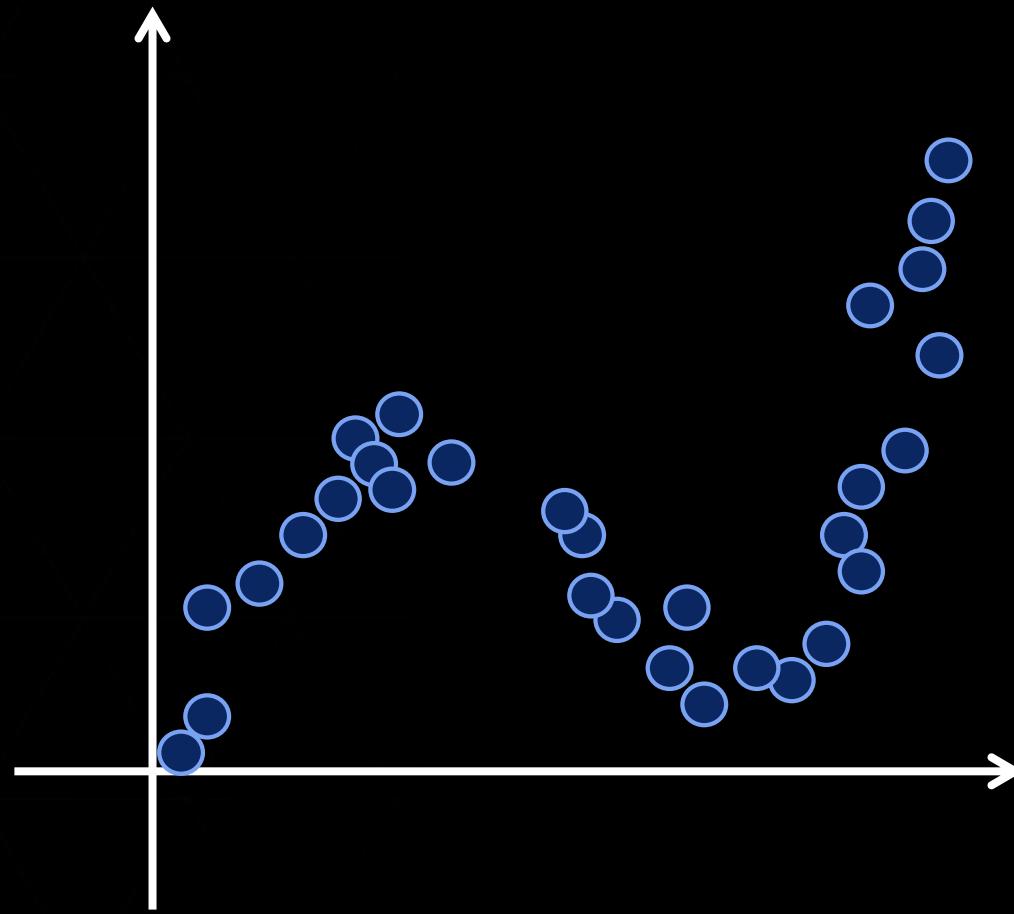


ОБУЧЕНИЕ С УЧИТЕЛЕМ: ЗАДАЧИ

Классификация

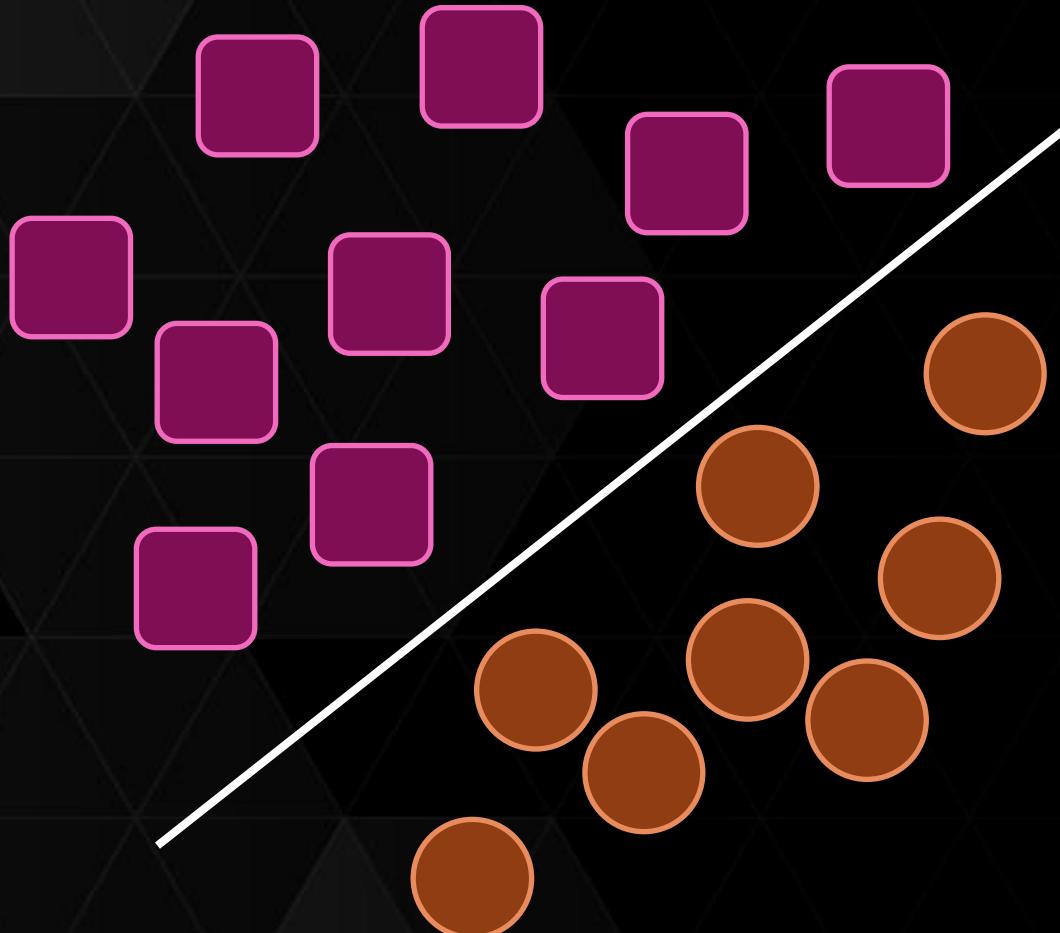


Регрессия

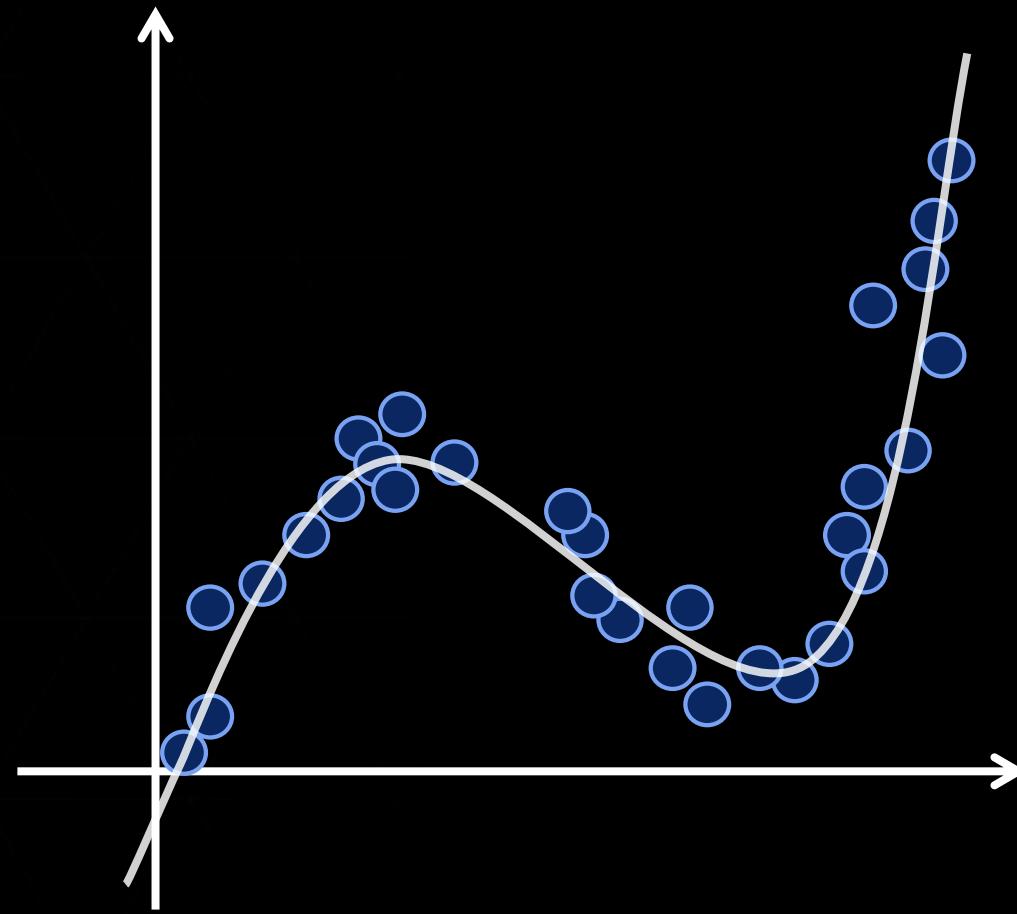


ОБУЧЕНИЕ С УЧИТЕЛЕМ: ЗАДАЧИ

Классификация



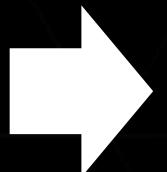
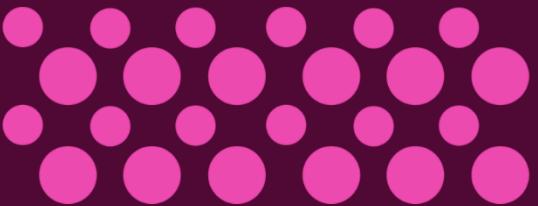
Регрессия



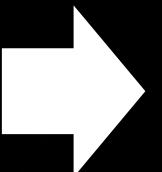
ТИПЫ ДАННЫХ И ЗАДАЧИ

Вход

Низкоразмерная
информация

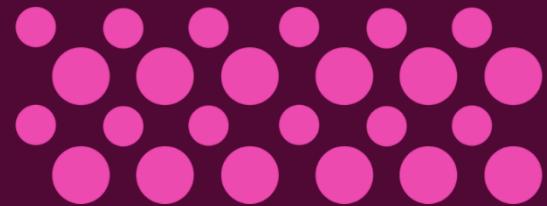


Модель



Выход

Низкоразмерная
информация

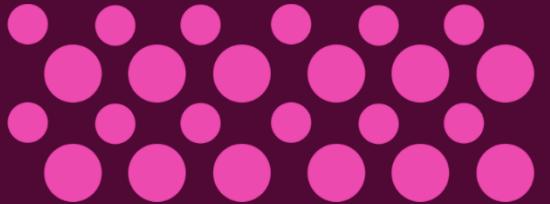


ТИПЫ ДАННЫХ И ЗАДАЧИ

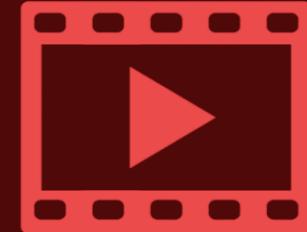


ТИПЫ ДАННЫХ И ЗАДАЧИ

Низкоразмерная
информация

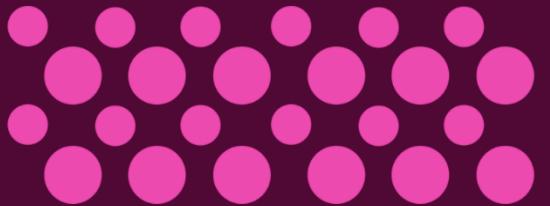


Изображение / видео



ТИПЫ ДАННЫХ И ЗАДАЧИ

Низкоразмерная
информация



Изображение / видео

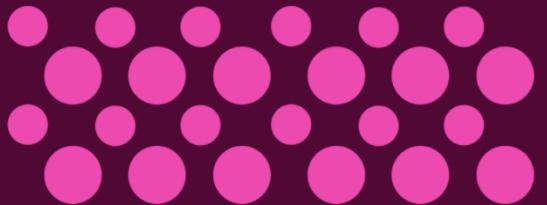


Текст



ТИПЫ ДАННЫХ И ЗАДАЧИ

Низкоразмерная
информация



Изображение / видео



Текст



Аудио



ТИПЫ ДАННЫХ И ЗАДАЧИ



РЕШАЕМЫЕ ЗАДАЧИ

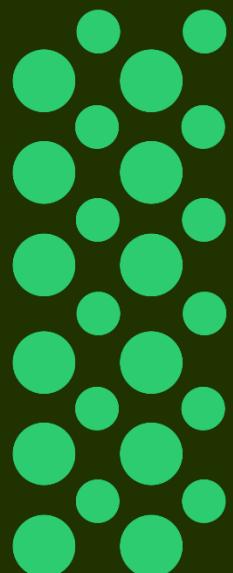
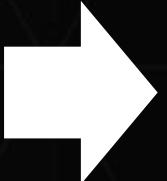
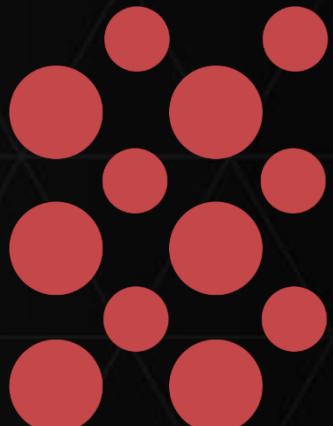
DEEP LEARNING



ОБУЧЕНИЕ ПРЕДСТАВЛЕНИЯ

Модель

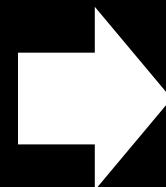
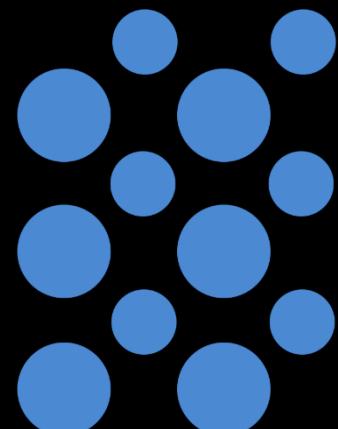
Вход



Промежуточное
представление

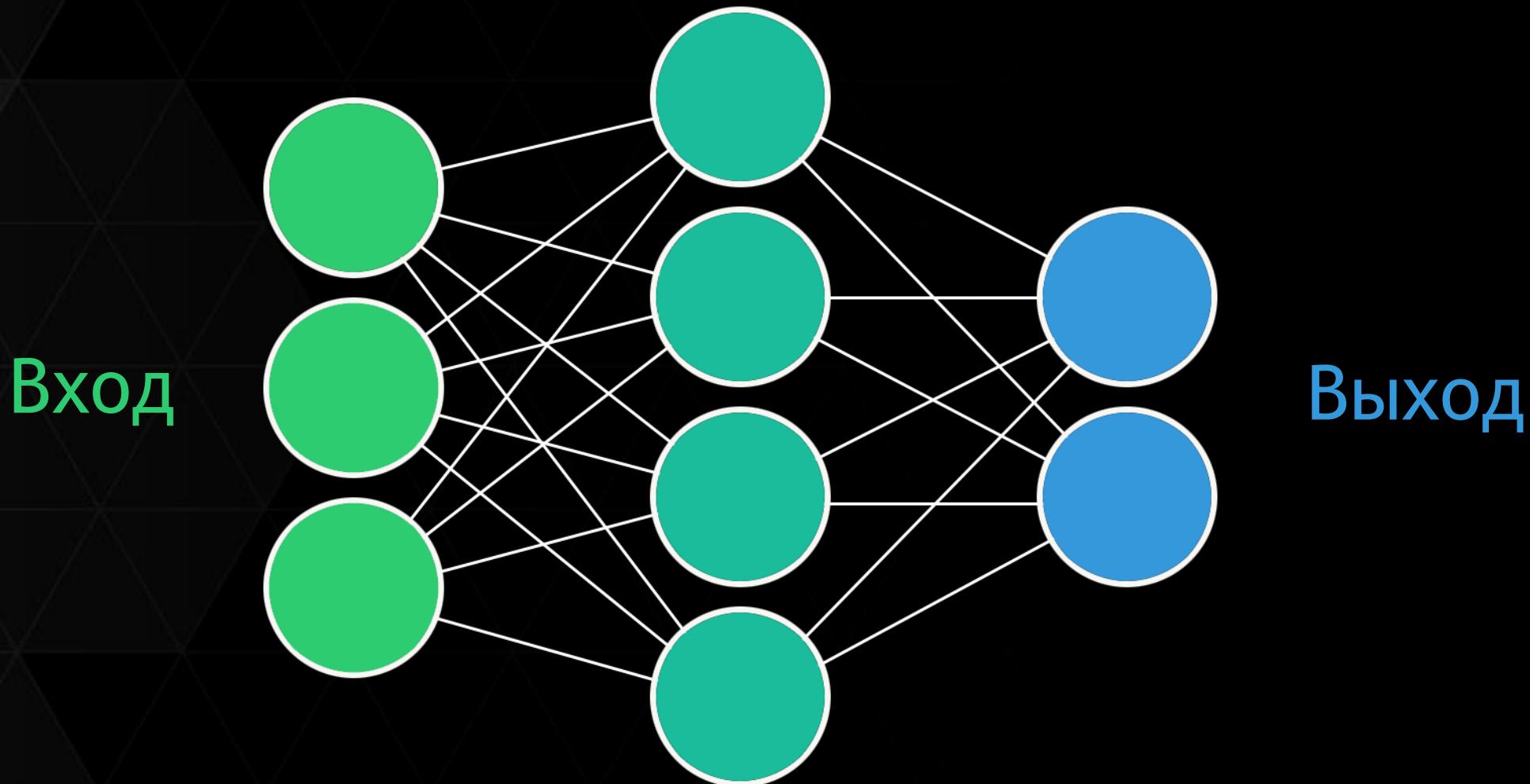


Выход

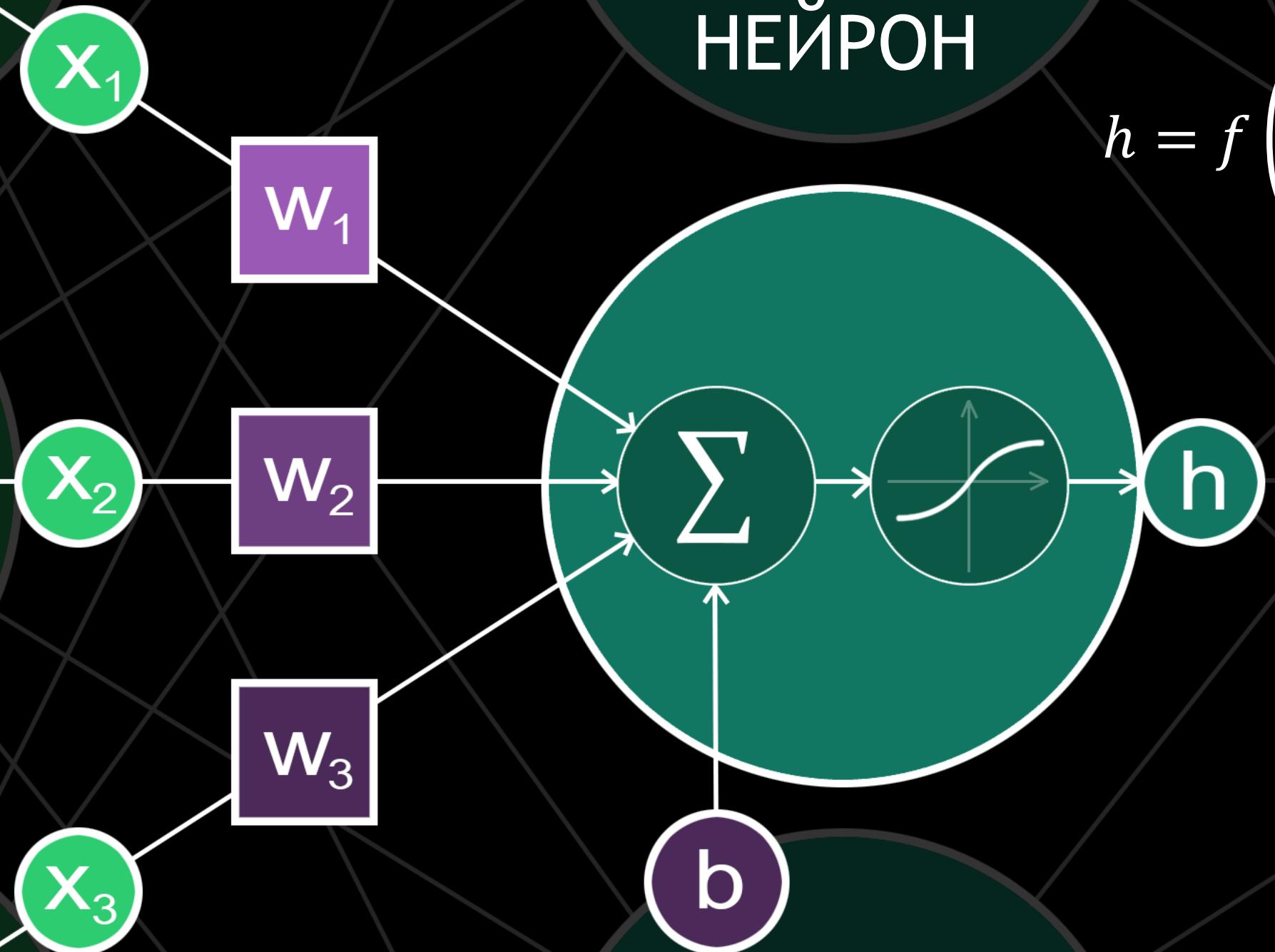


НЕЙРОННЫЕ СЕТИ

НЕЙРОННЫЕ СЕТИ

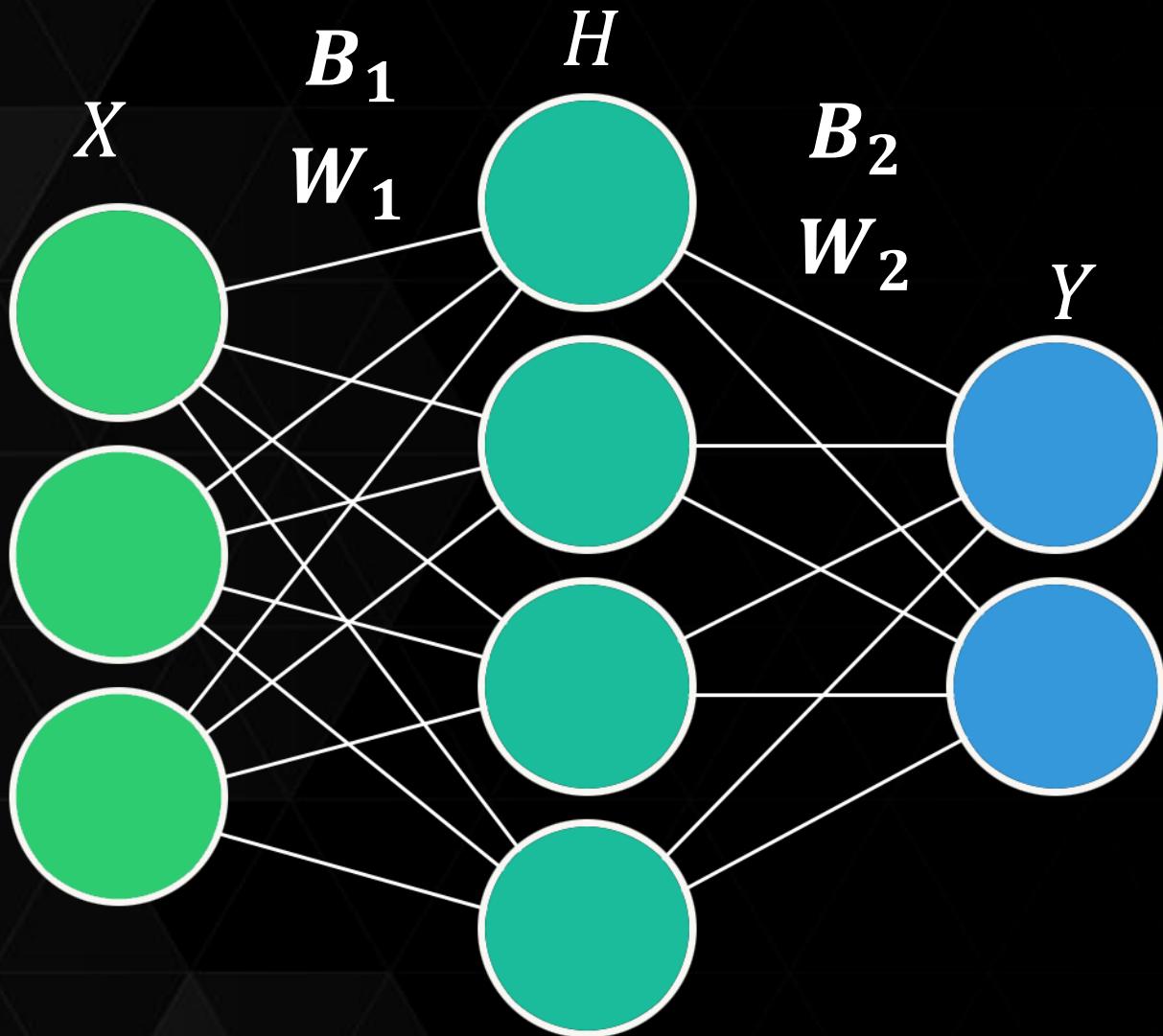


НЕЙРОН



$$h = f \left(\sum_i w_i x_i + b \right)$$

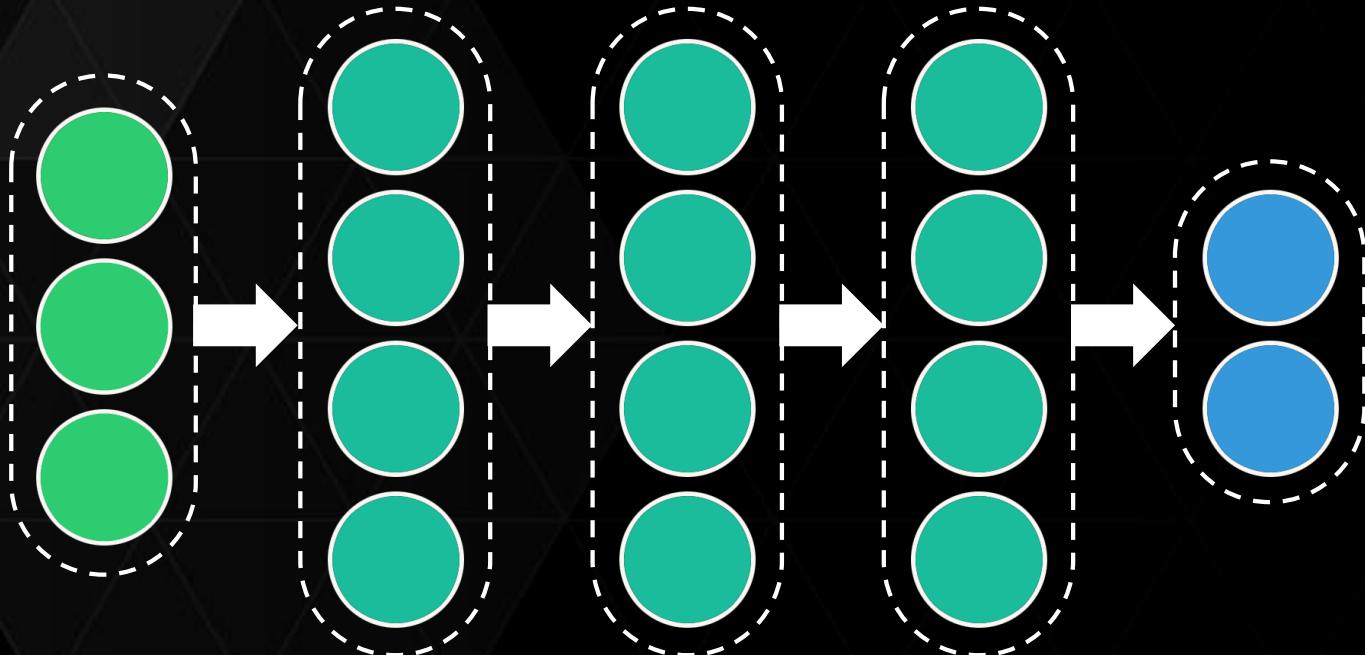
НЕЙРОННЫЕ СЕТИ



$$H = f(W_1X + B_1)$$

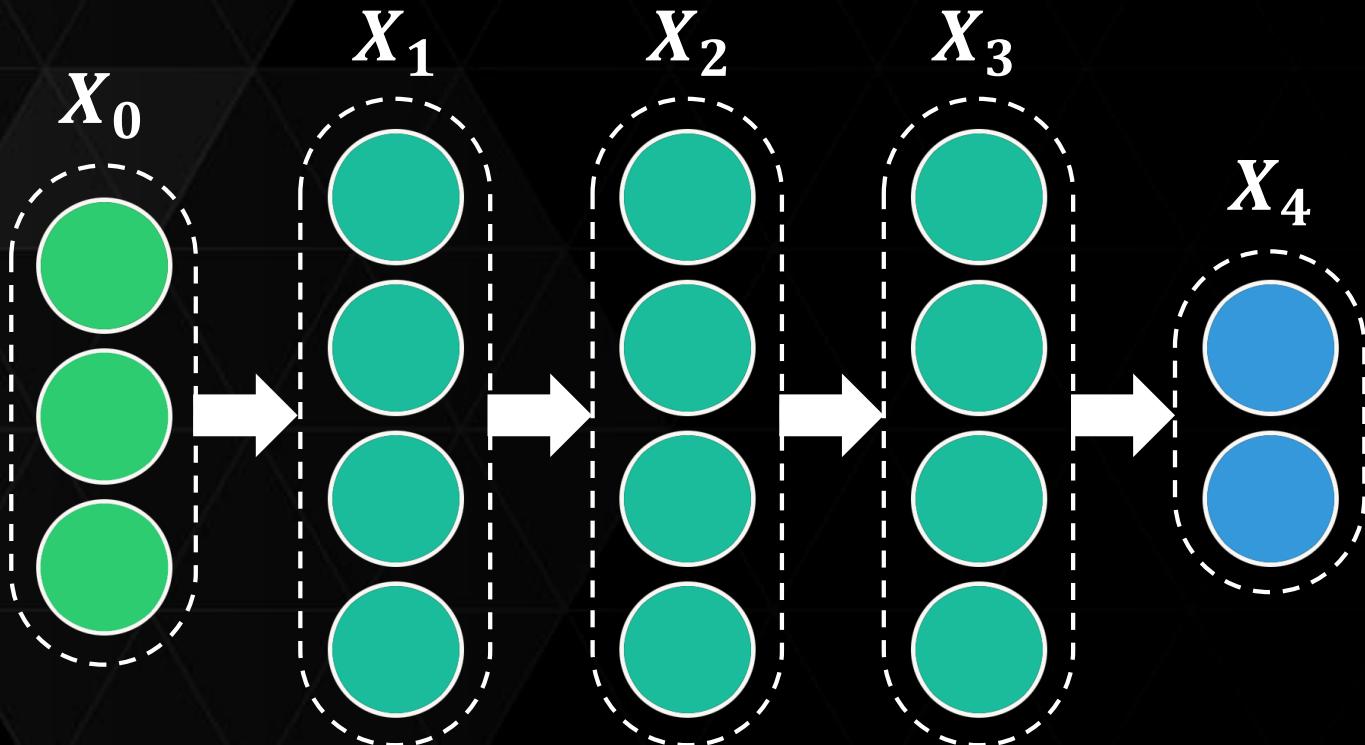
$$Y = f(W_2H + B_2)$$

МНОГОСЛОЙНЫЕ НЕЙРОННЫЕ СЕТИ



Промежуточное
представление

МНОГОСЛОЙНЫЕ НЕЙРОННЫЕ СЕТИ



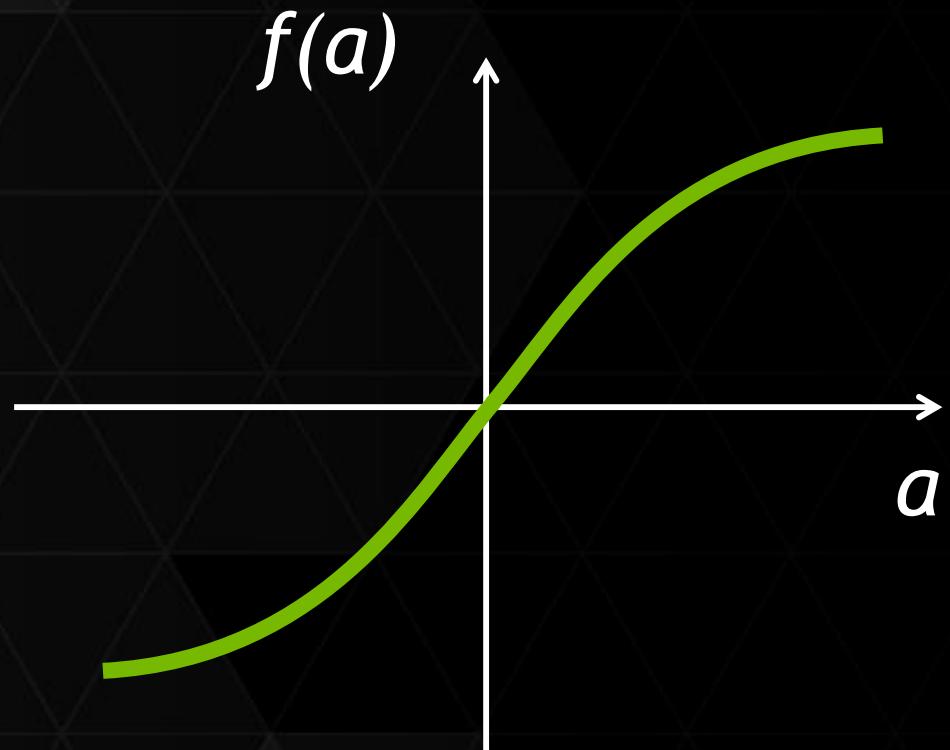
Промежуточное
представление

$$X_k = f(W_k X_{k-1} + B_k)$$

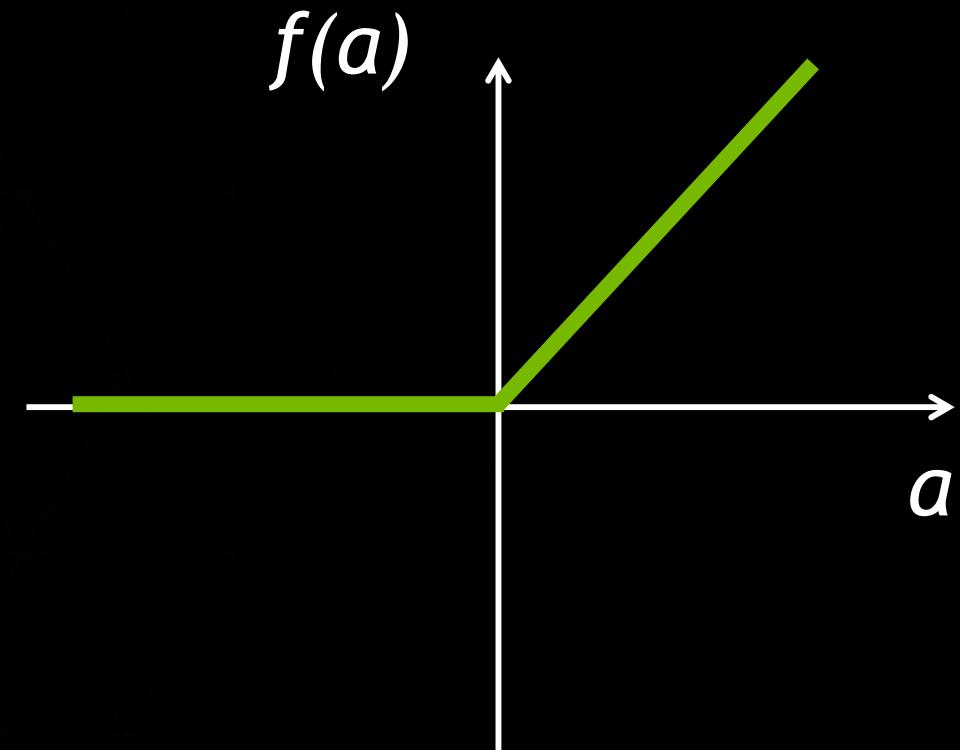
$$X_N = G_W(X_0)$$

ФУНКЦИЯ АКТИВАЦИИ

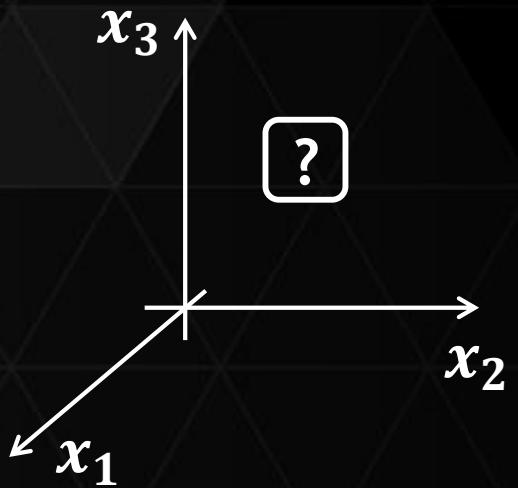
Sigmoid



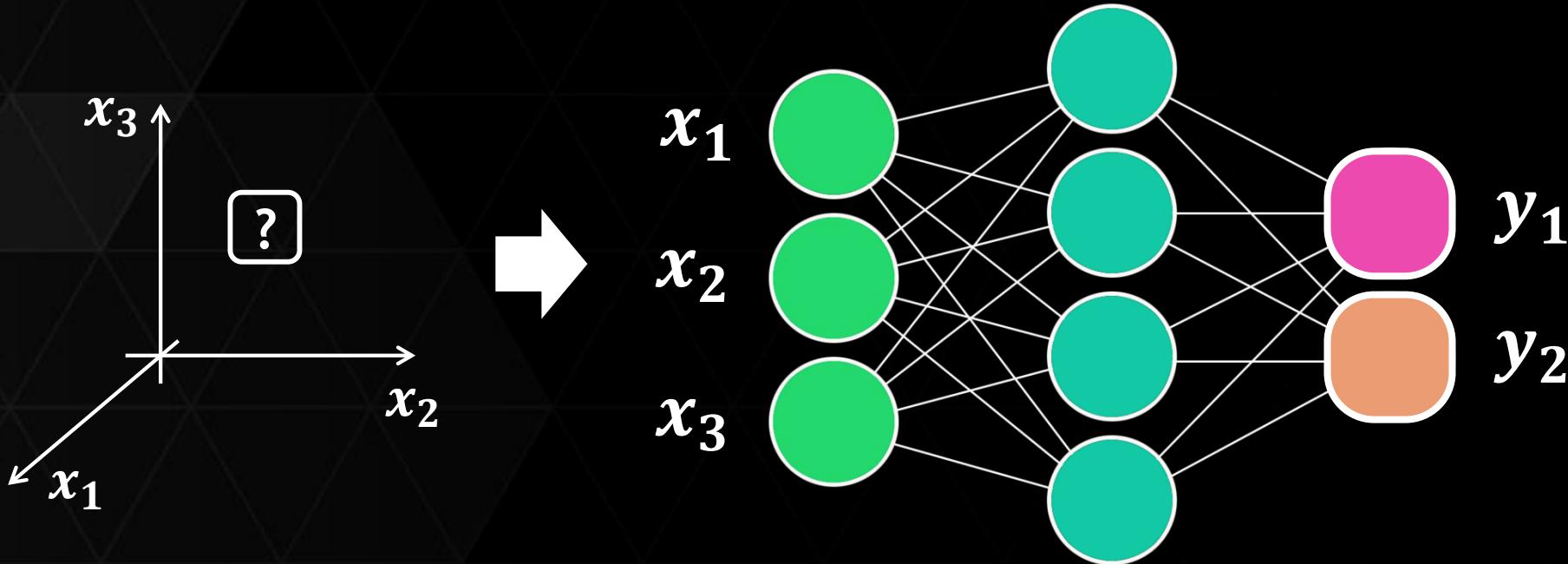
ReLU



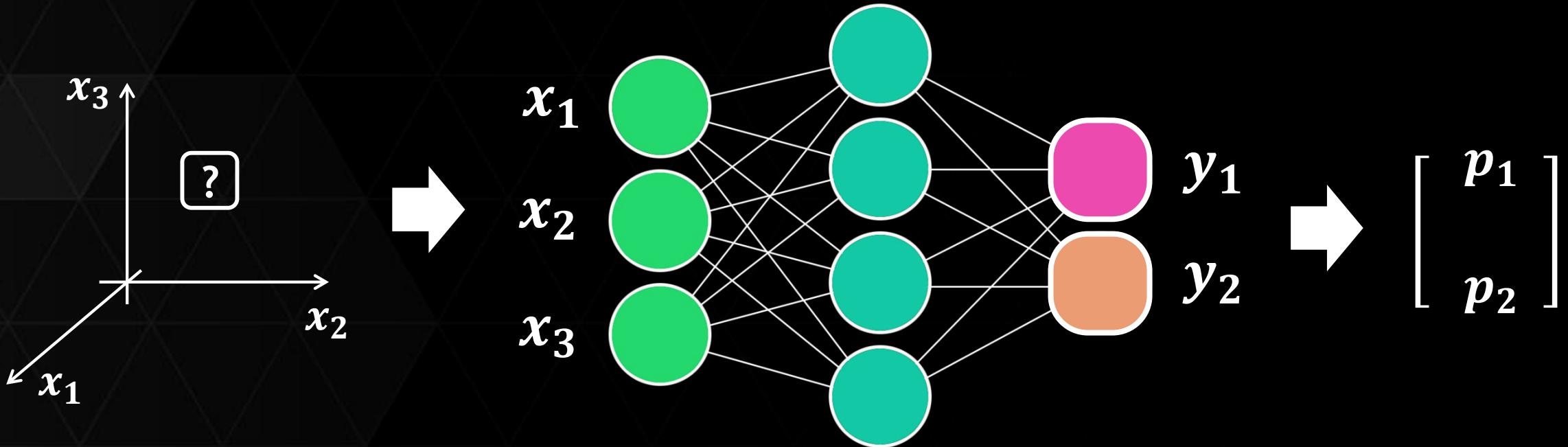
ПРИМЕР: КЛАССИФИКАЦИЯ ОБЪЕКТА



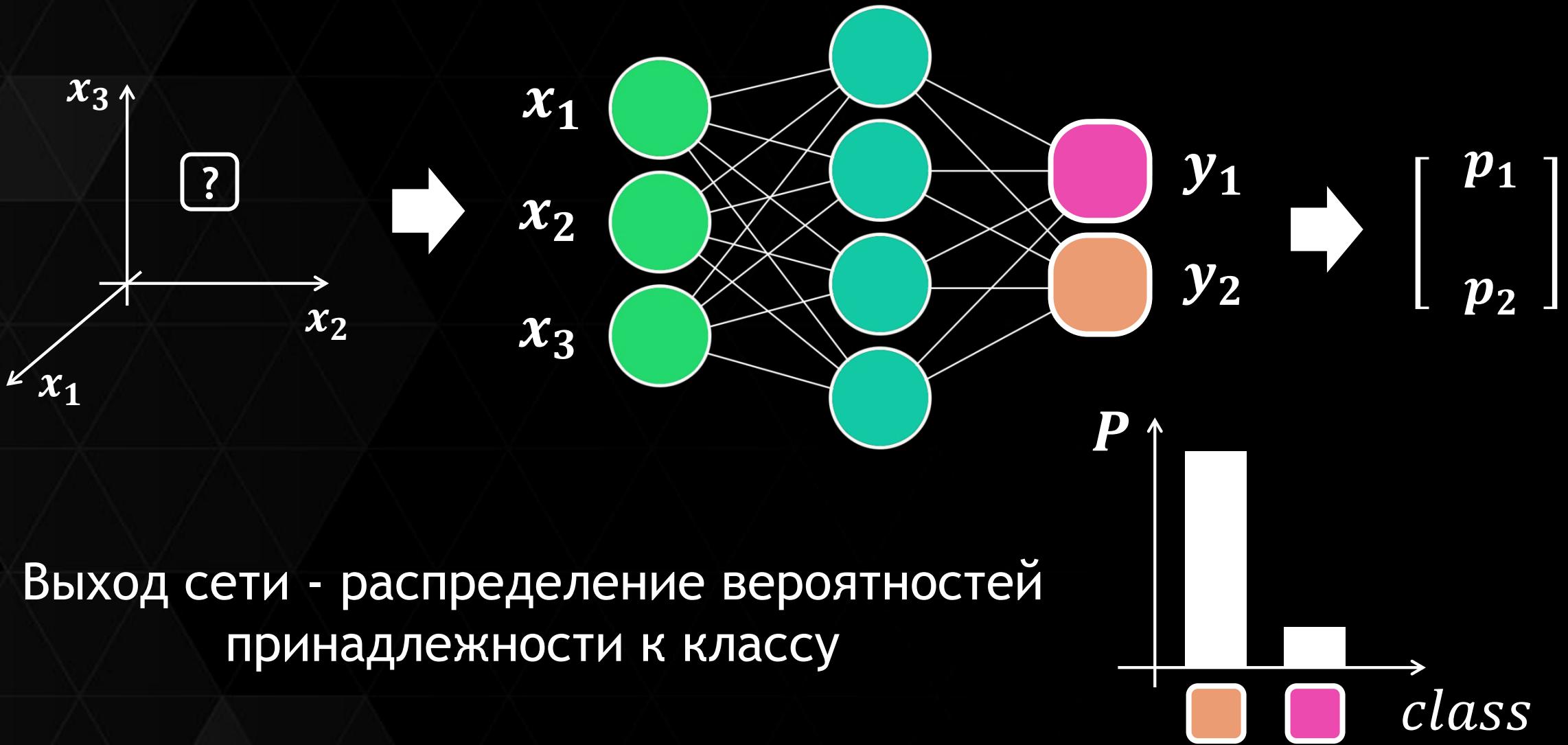
ПРИМЕР: КЛАССИФИКАЦИЯ ОБЪЕКТА



ПРИМЕР: КЛАССИФИКАЦИЯ ОБЪЕКТА



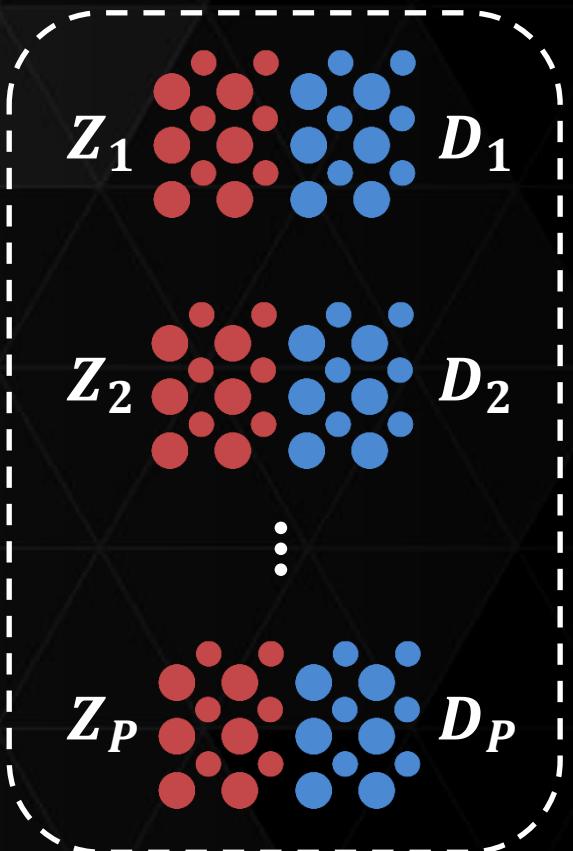
ПРИМЕР: КЛАССИФИКАЦИЯ ОБЪЕКТА



ОБУЧЕНИЕ НЕЙРОННЫХ СЕТЕЙ

НЕЙРОННЫЕ СЕТИ: ОБУЧЕНИЕ

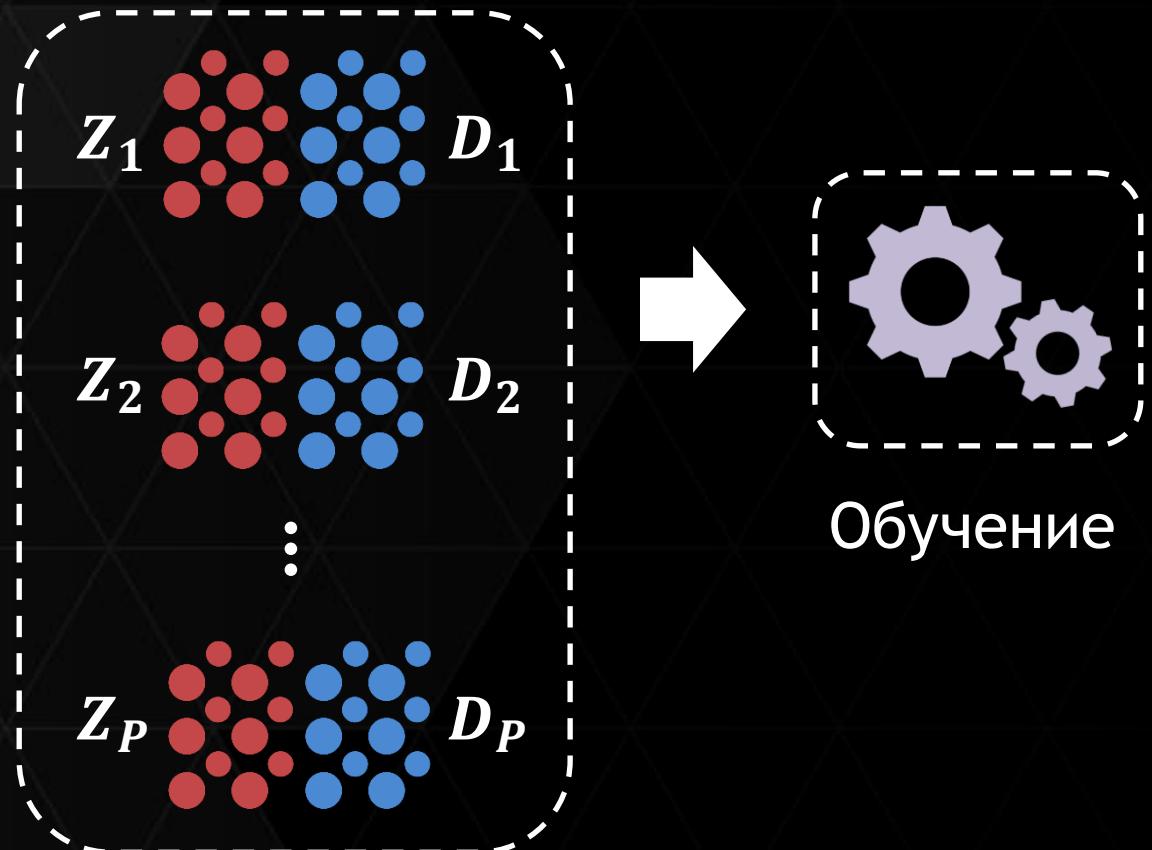
Обучение с учителем (на примерах)



Обучающая выборка

НЕЙРОННЫЕ СЕТИ: ОБУЧЕНИЕ

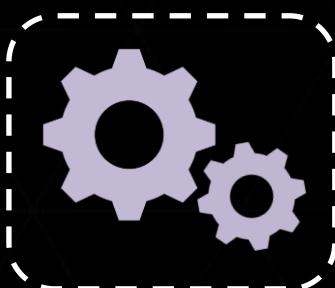
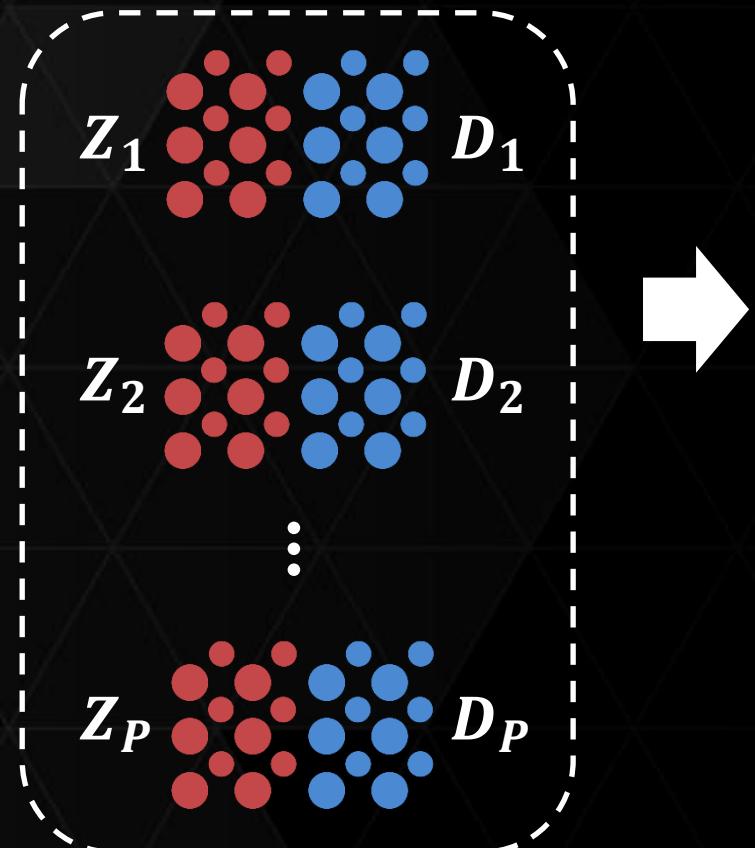
Обучение с учителем (на примерах)



Обучение

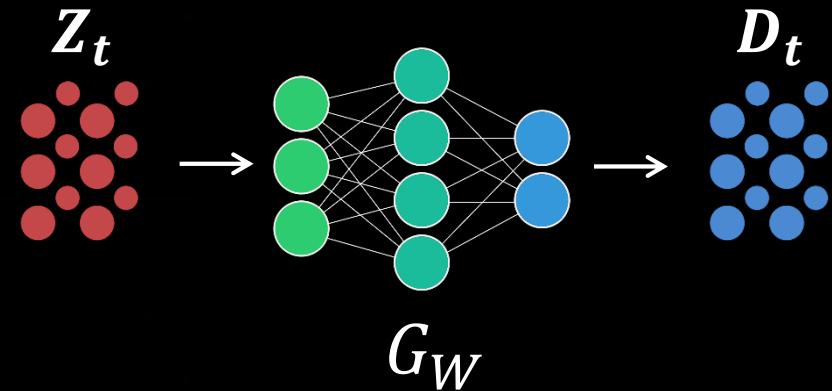
НЕЙРОННЫЕ СЕТИ: ОБУЧЕНИЕ

Обучение с учителем (на примерах)



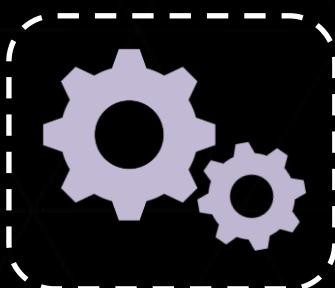
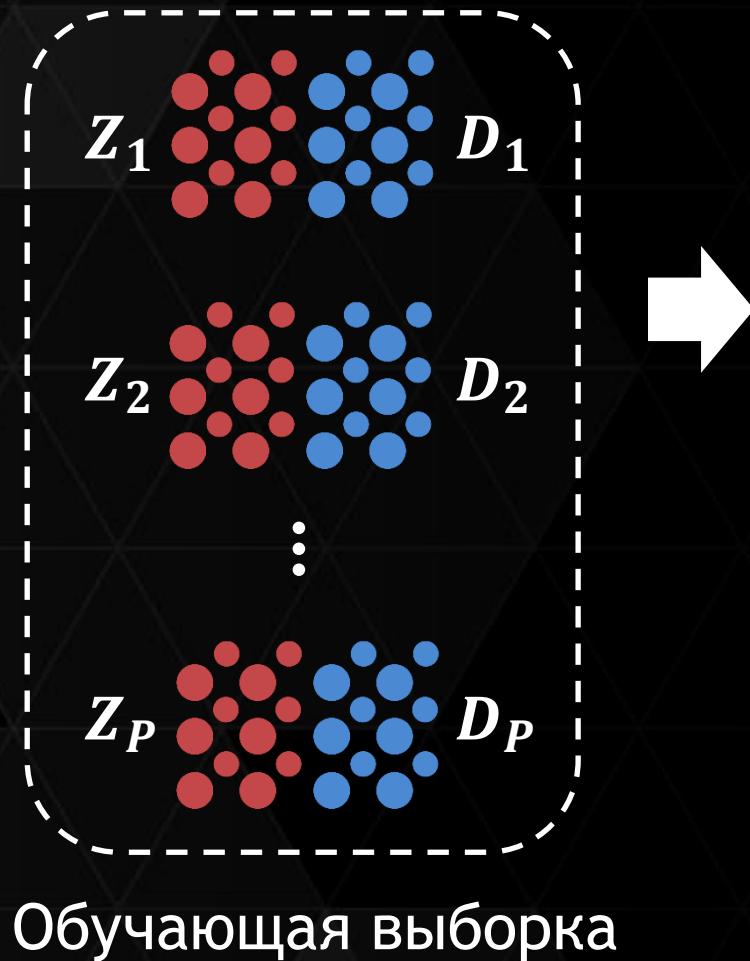
Обучающая выборка

Найти W :



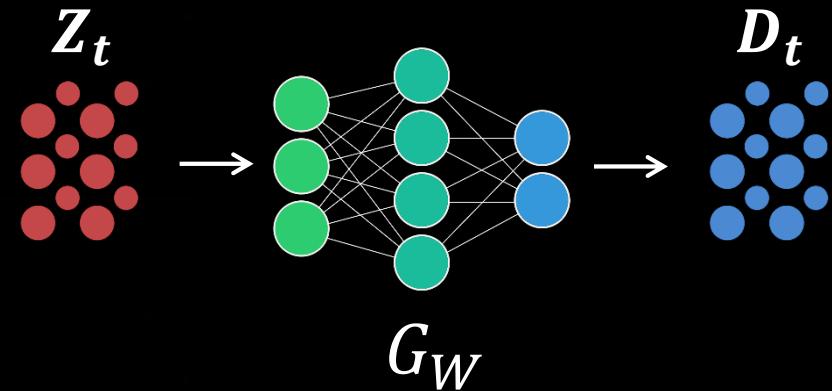
НЕЙРОННЫЕ СЕТИ: ОБУЧЕНИЕ

Обучение с учителем (на примерах)



Обучение

Найти W :

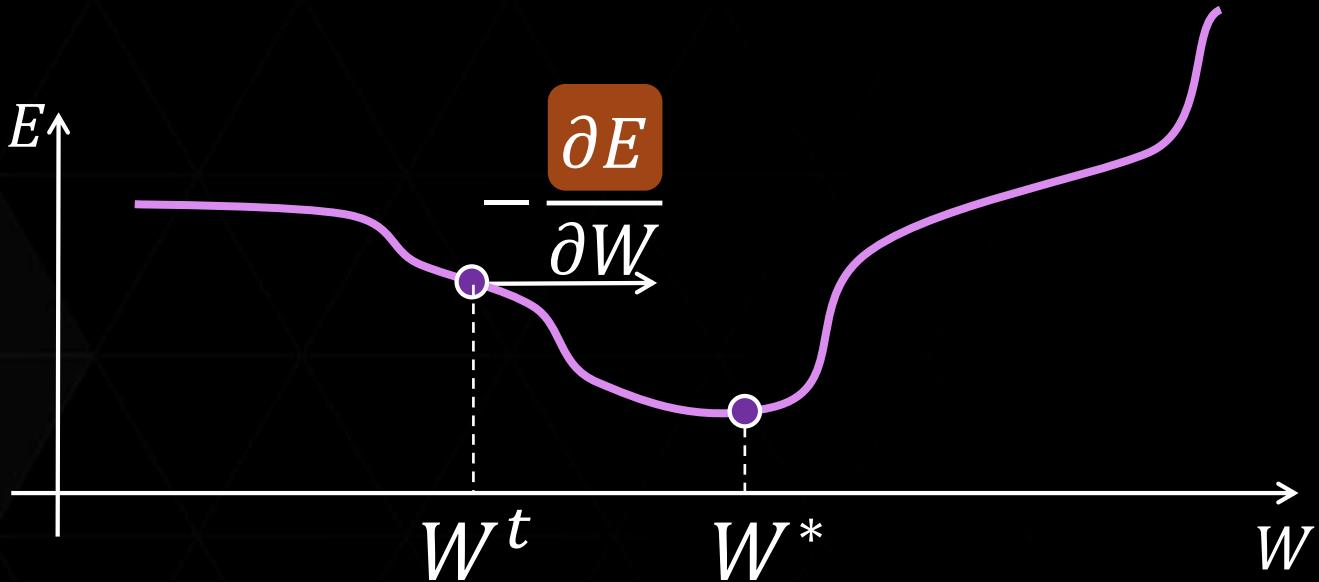


$$E = \| D - G_W(Z) \|^2$$

$$W^* = \operatorname{argmin}_W E(W)$$

НЕЙРОННЫЕ СЕТИ: ОБУЧЕНИЕ

Градиентный спуск



Итерационный процесс

$$W^{t+1} = W^t - \alpha \frac{\partial E}{\partial W}(W^t)$$

НЕЙРОННЫЕ СЕТИ: ОБУЧЕНИЕ

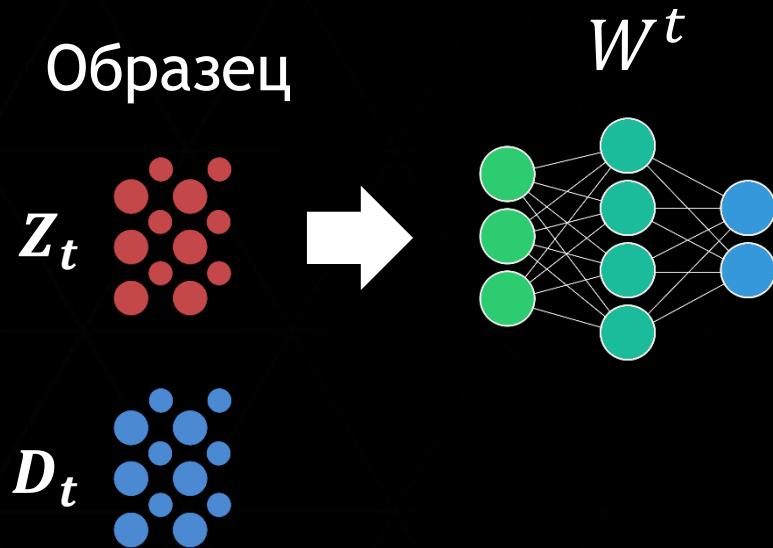
Образец

На итерации t :



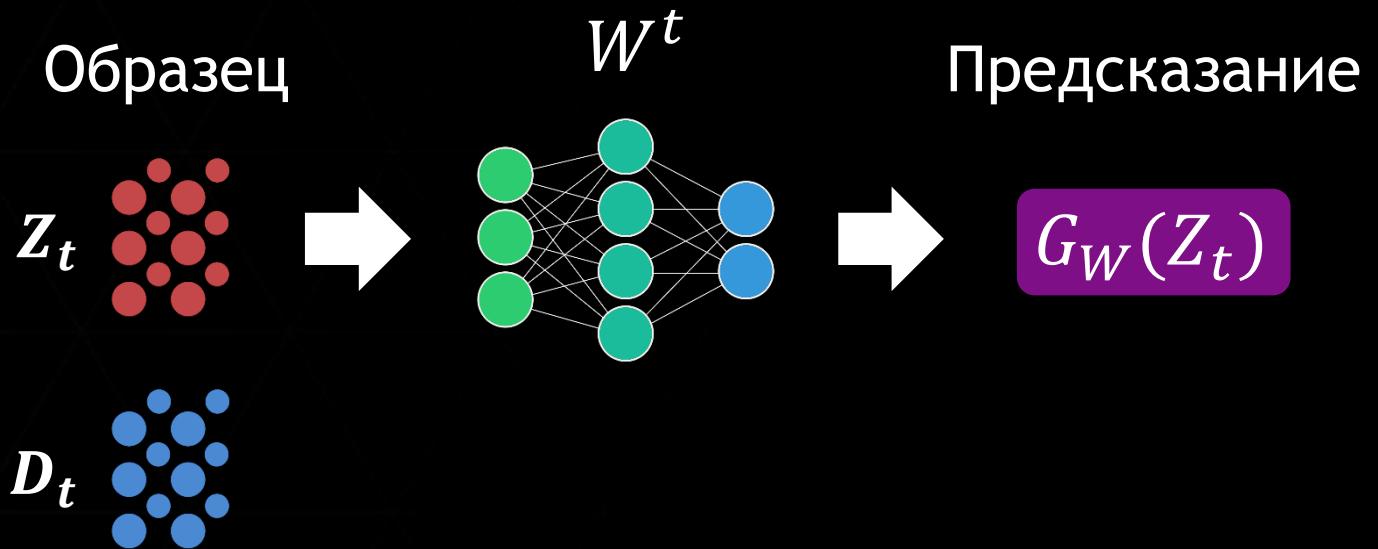
НЕЙРОННЫЕ СЕТИ: ОБУЧЕНИЕ

На итерации t :



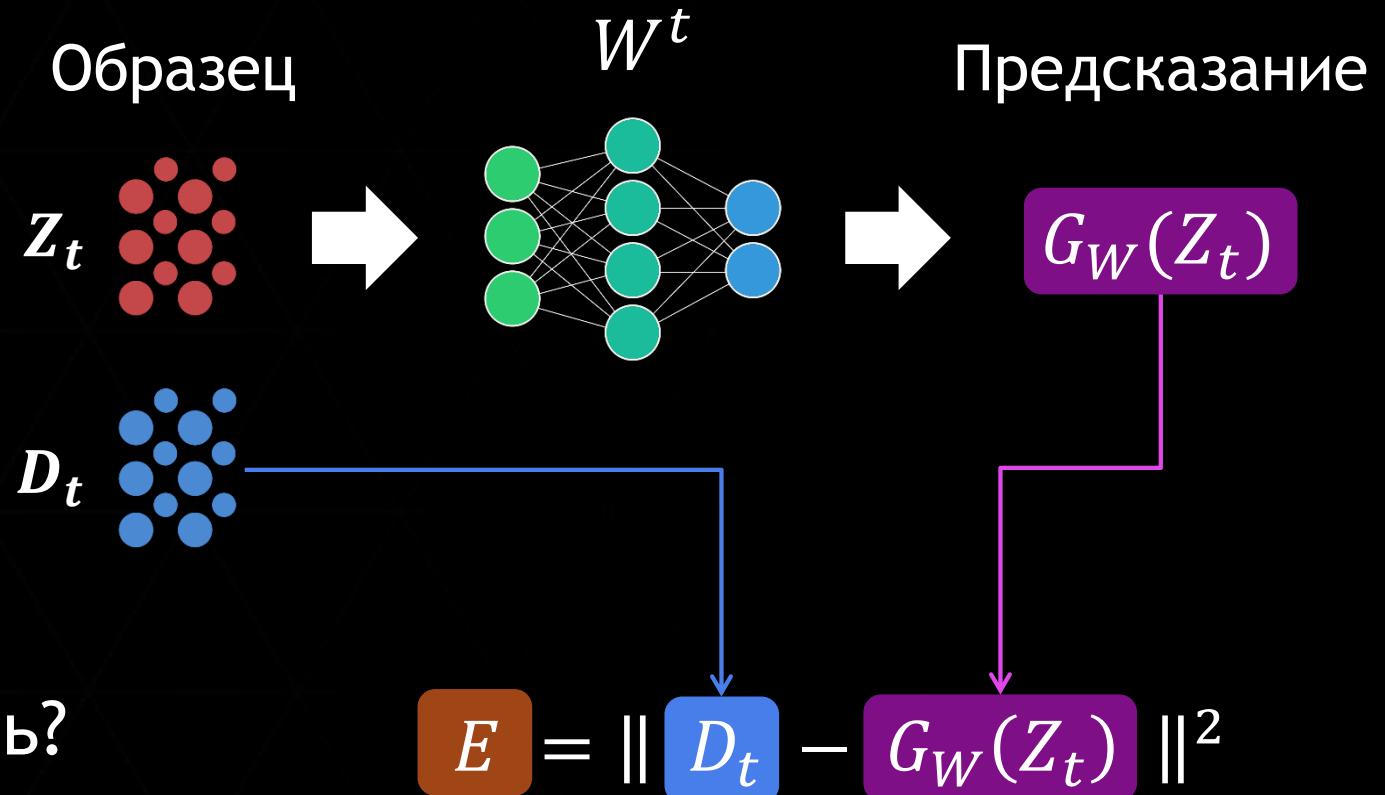
НЕЙРОННЫЕ СЕТИ: ОБУЧЕНИЕ

На итерации t :



НЕЙРОННЫЕ СЕТИ: ОБУЧЕНИЕ

На итерации t :



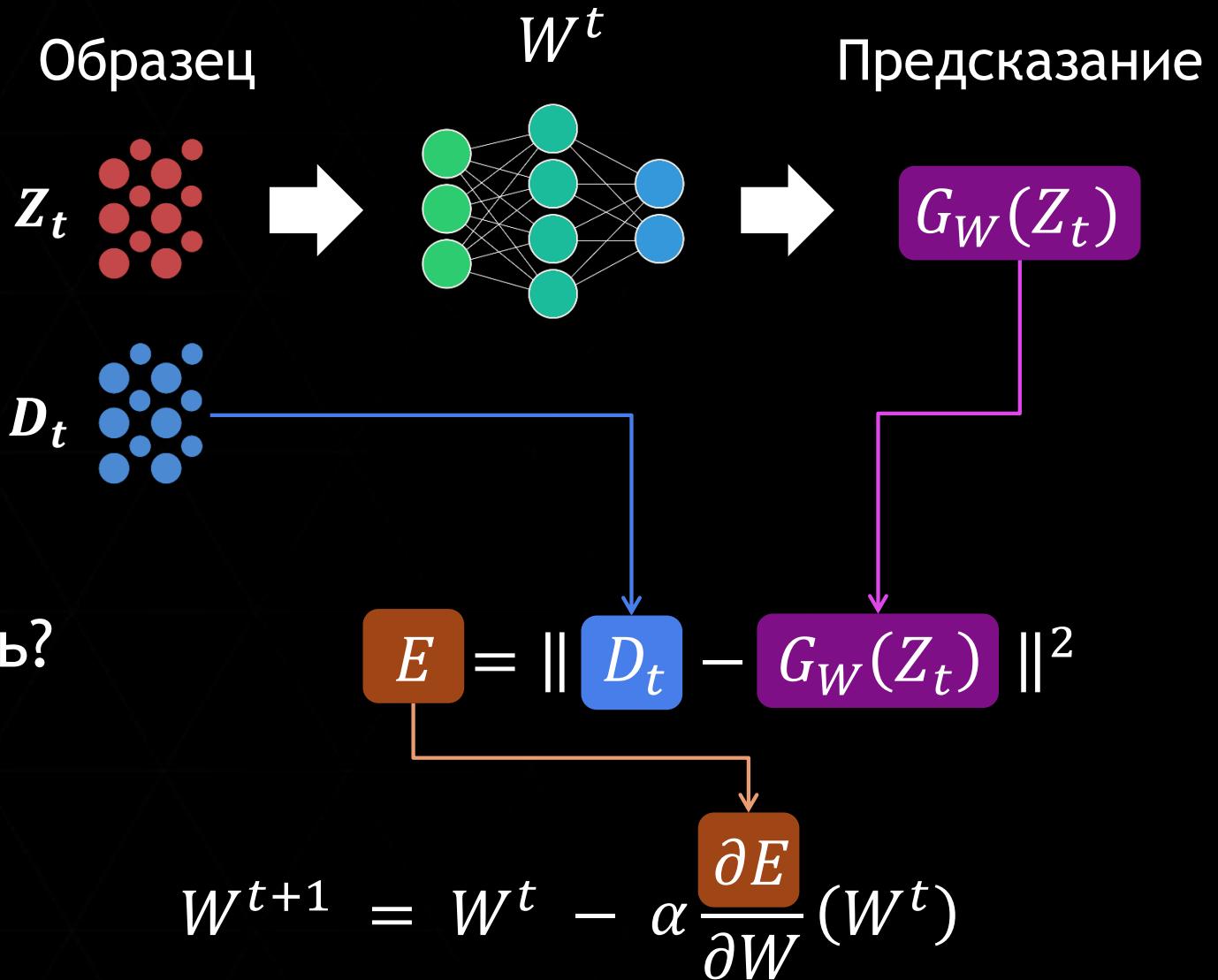
Как сильно мы ошиблись?

НЕЙРОННЫЕ СЕТИ: ОБУЧЕНИЕ

На итерации t :

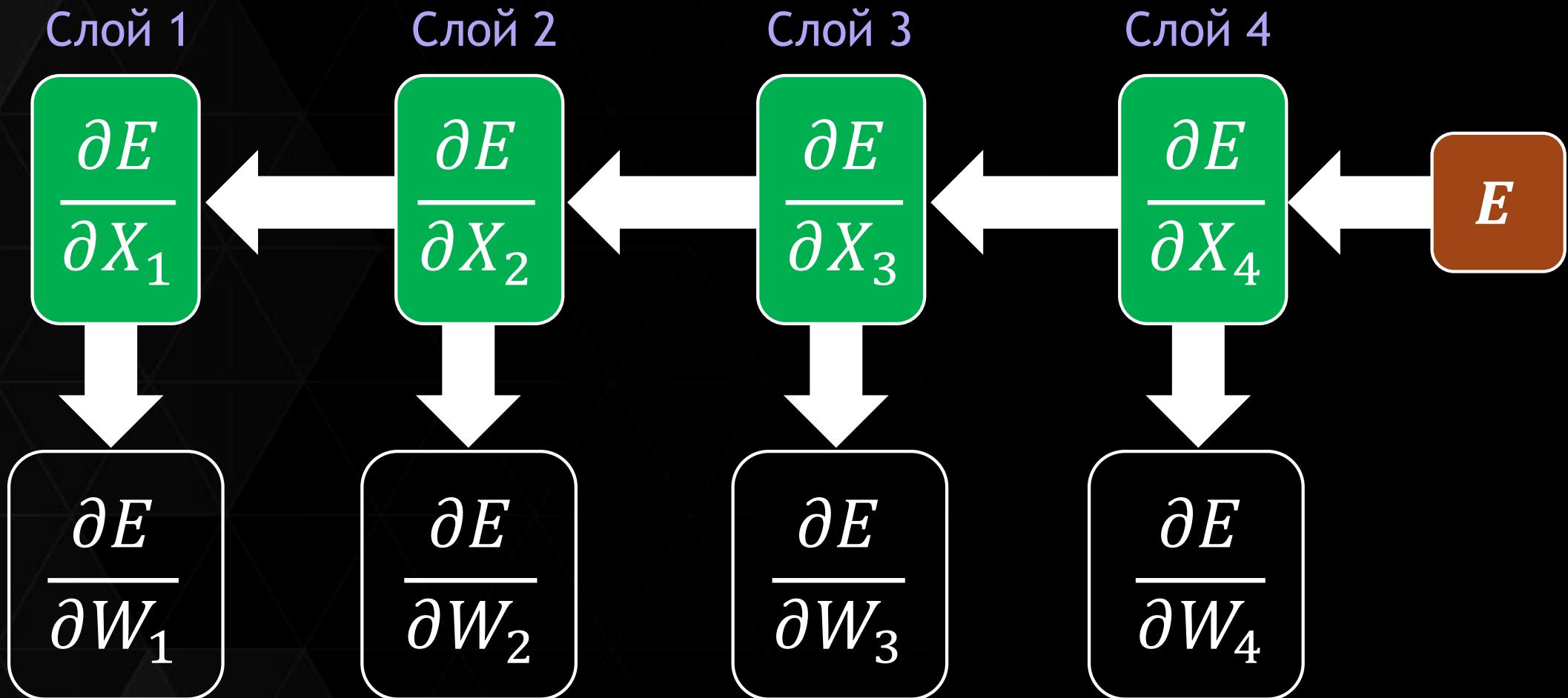
Как сильно мы ошиблись?

Поправить веса!



НЕЙРОННЫЕ СЕТИ: ОБУЧЕНИЕ

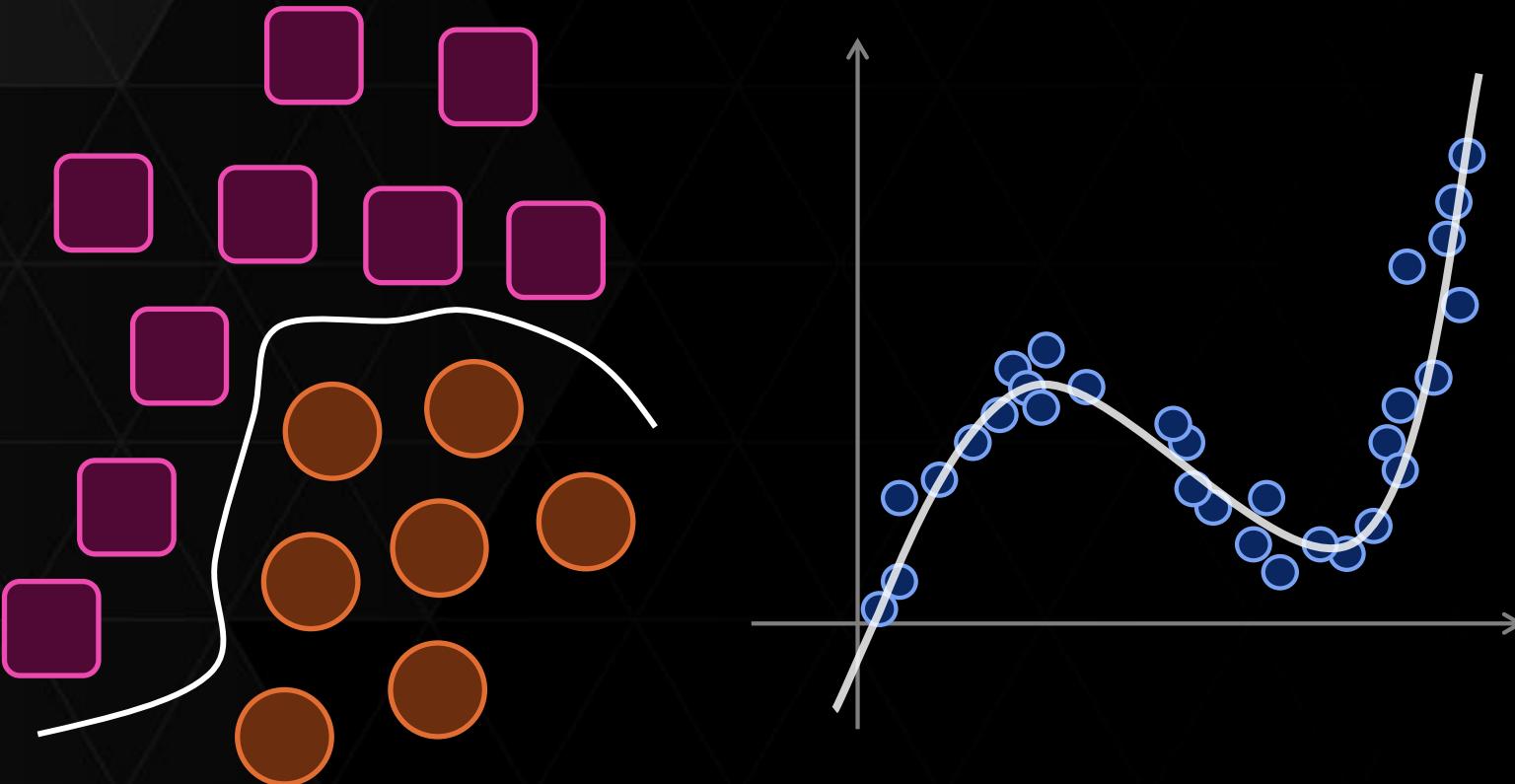
Обратное распространение ошибки



РЕШАЕМЫЕ ЗАДАЧИ

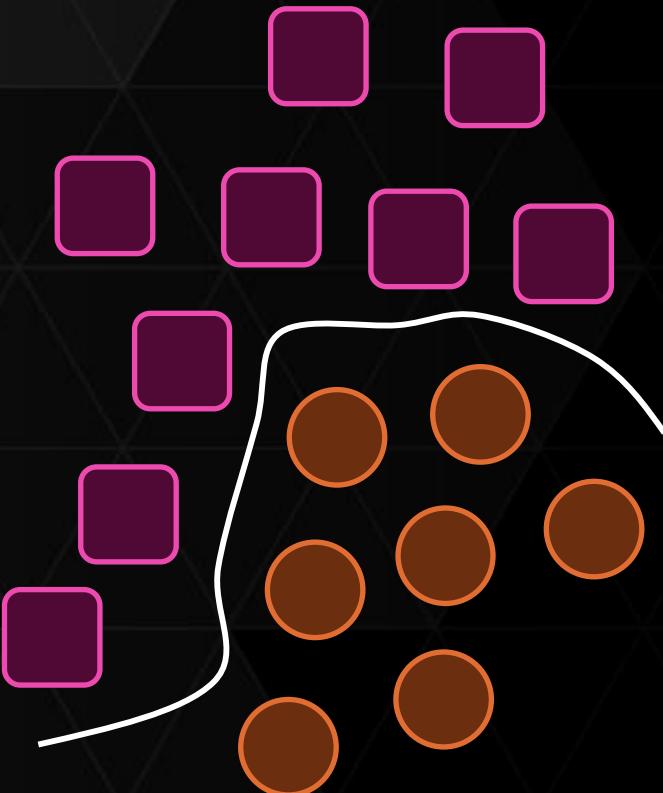
Классификация

Регрессия

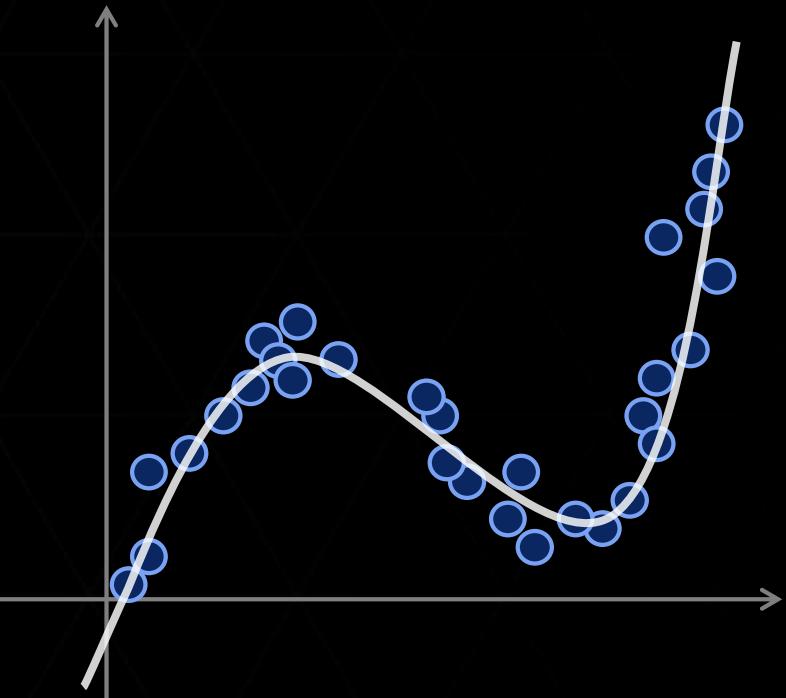


РЕШАЕМЫЕ ЗАДАЧИ

Классификация



Регрессия



Машинное зрение



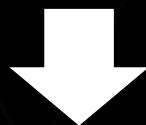
???



КОМПЬЮТЕРНОЕ ЗРЕНИЕ

КОМПЬЮТЕРНОЕ ЗРЕНИЕ: ЗАДАЧИ

Классификация



“Котик”

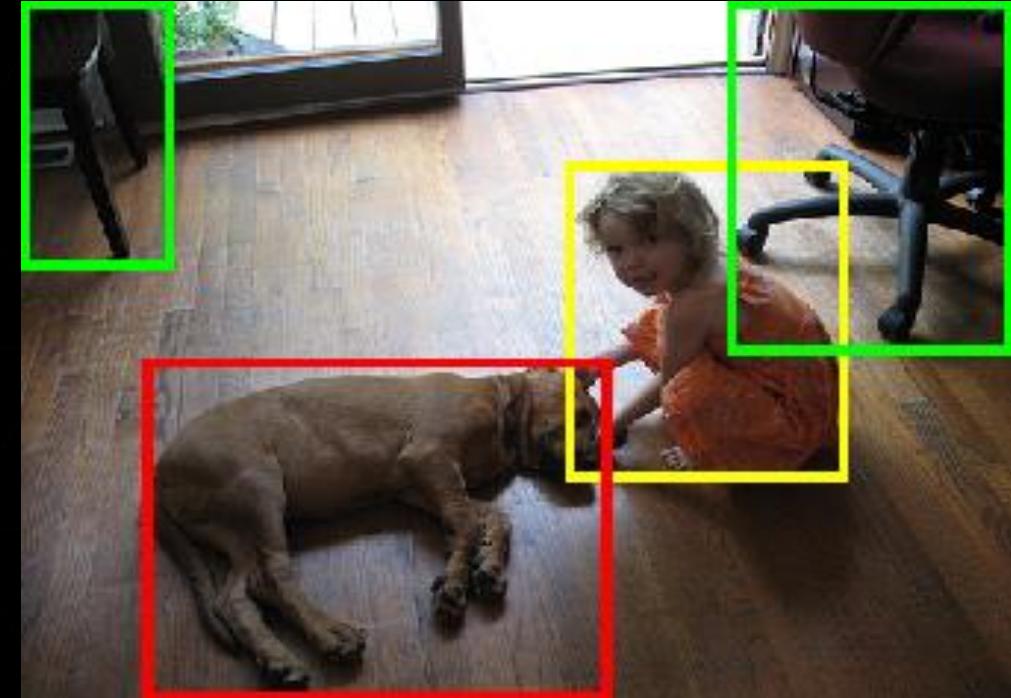
КОМПЬЮТЕРНОЕ ЗРЕНИЕ: ЗАДАЧИ

Классификация



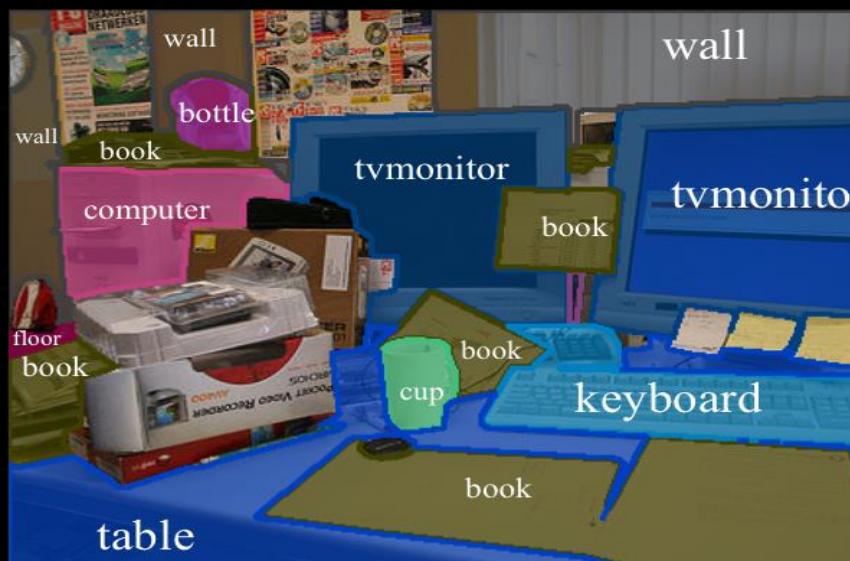
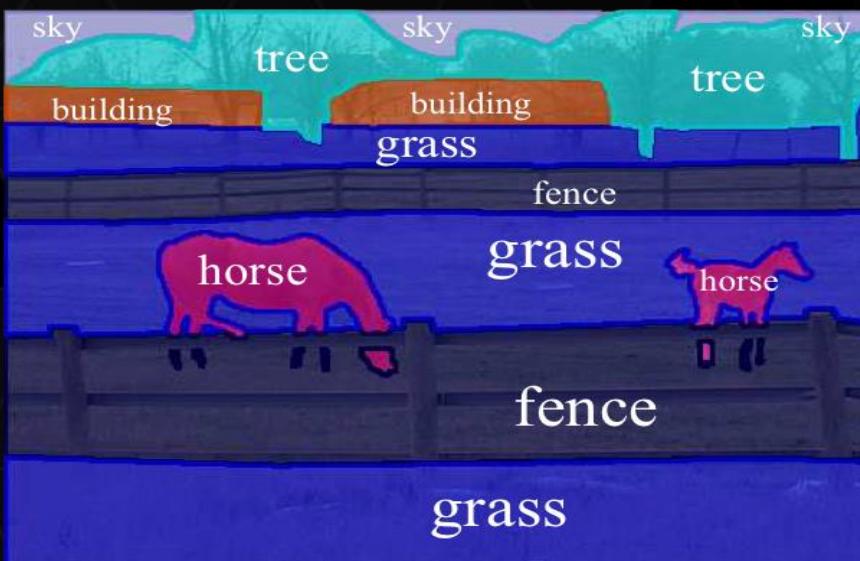
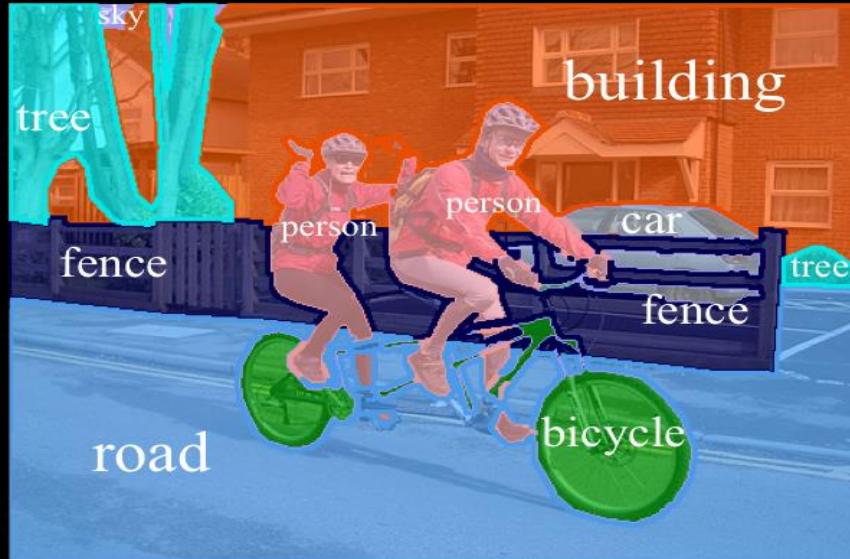
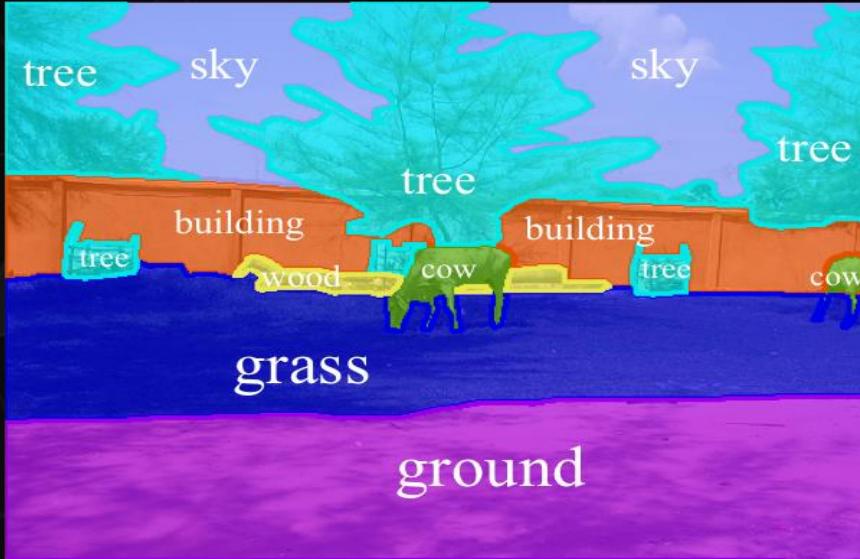
“Котик”

Детектирование объектов



- Человек
- Собака
- Стул

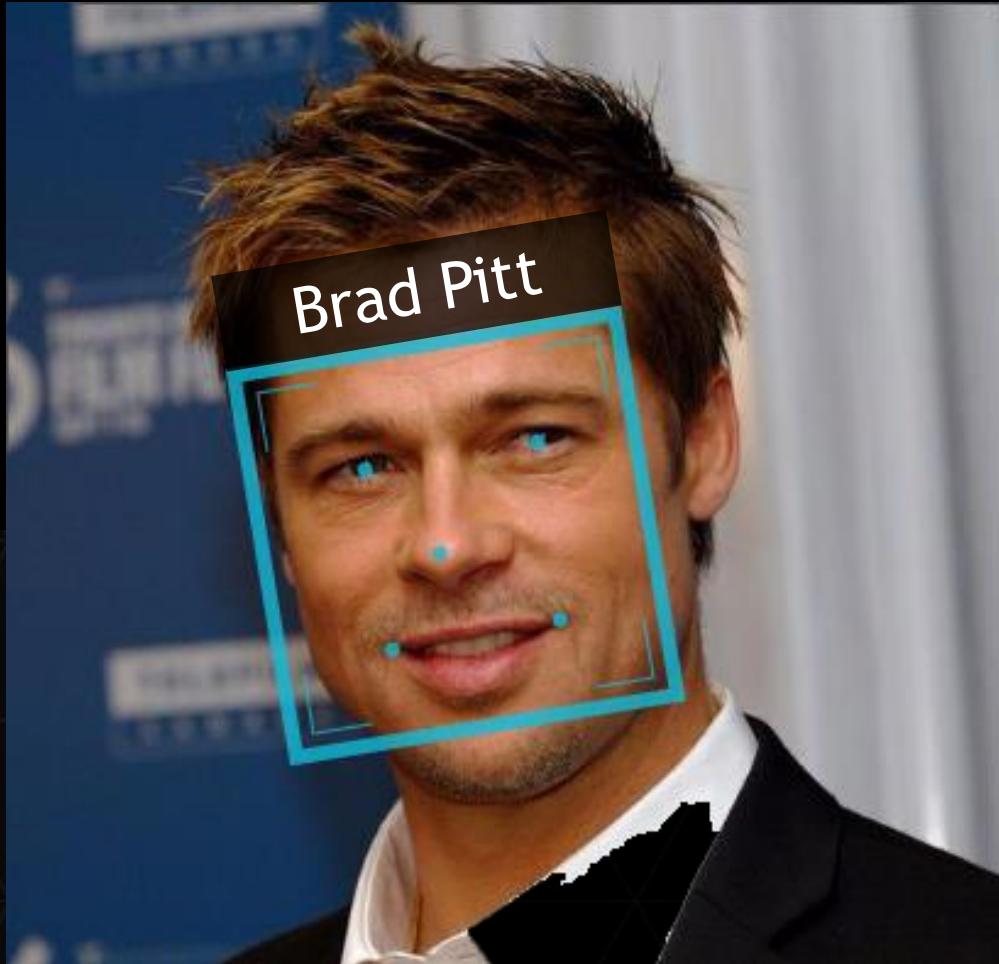
КОМПЬЮТЕРНОЕ ЗРЕНИЕ: ЗАДАЧИ



Семантическая
сегментация

РАСПОЗНАВАНИЕ ЛИЦ

Идентификация

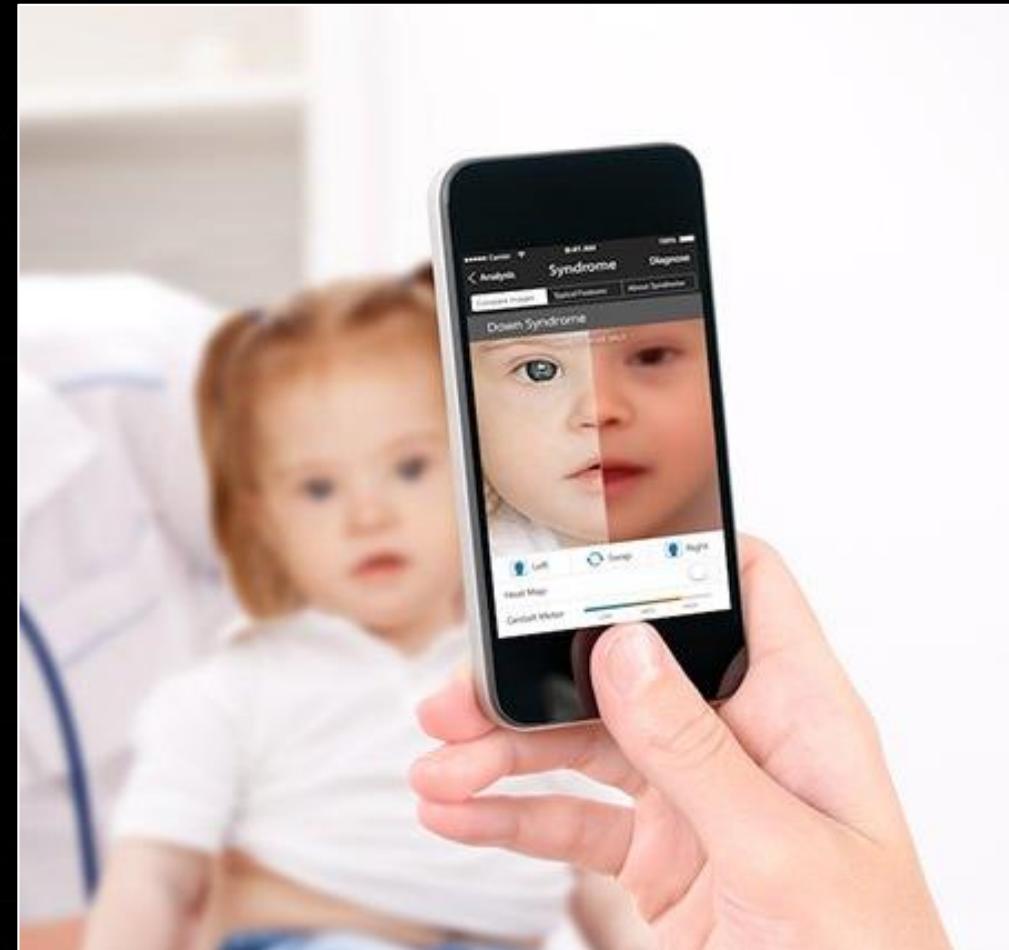
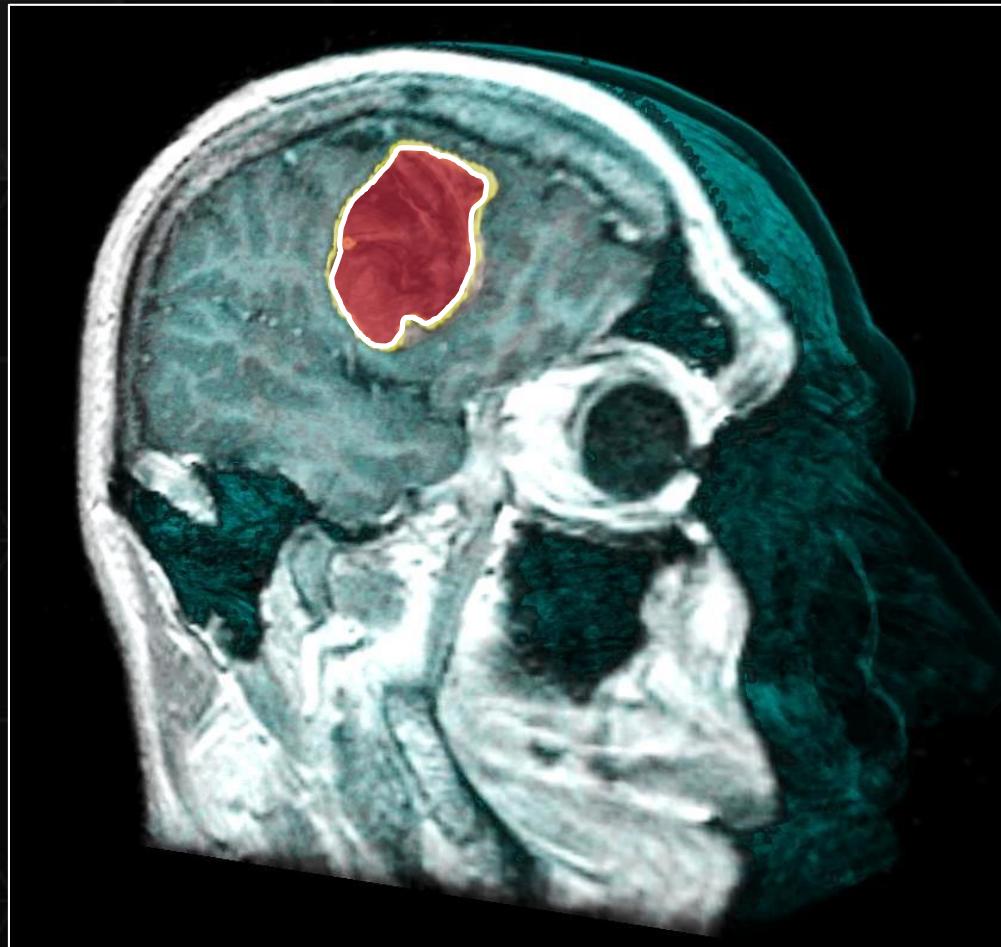


Пол, возраст, эмоции

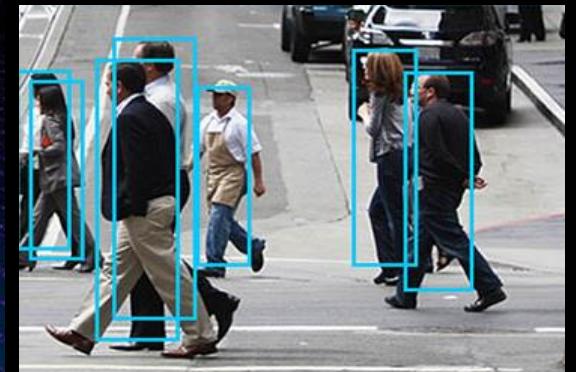
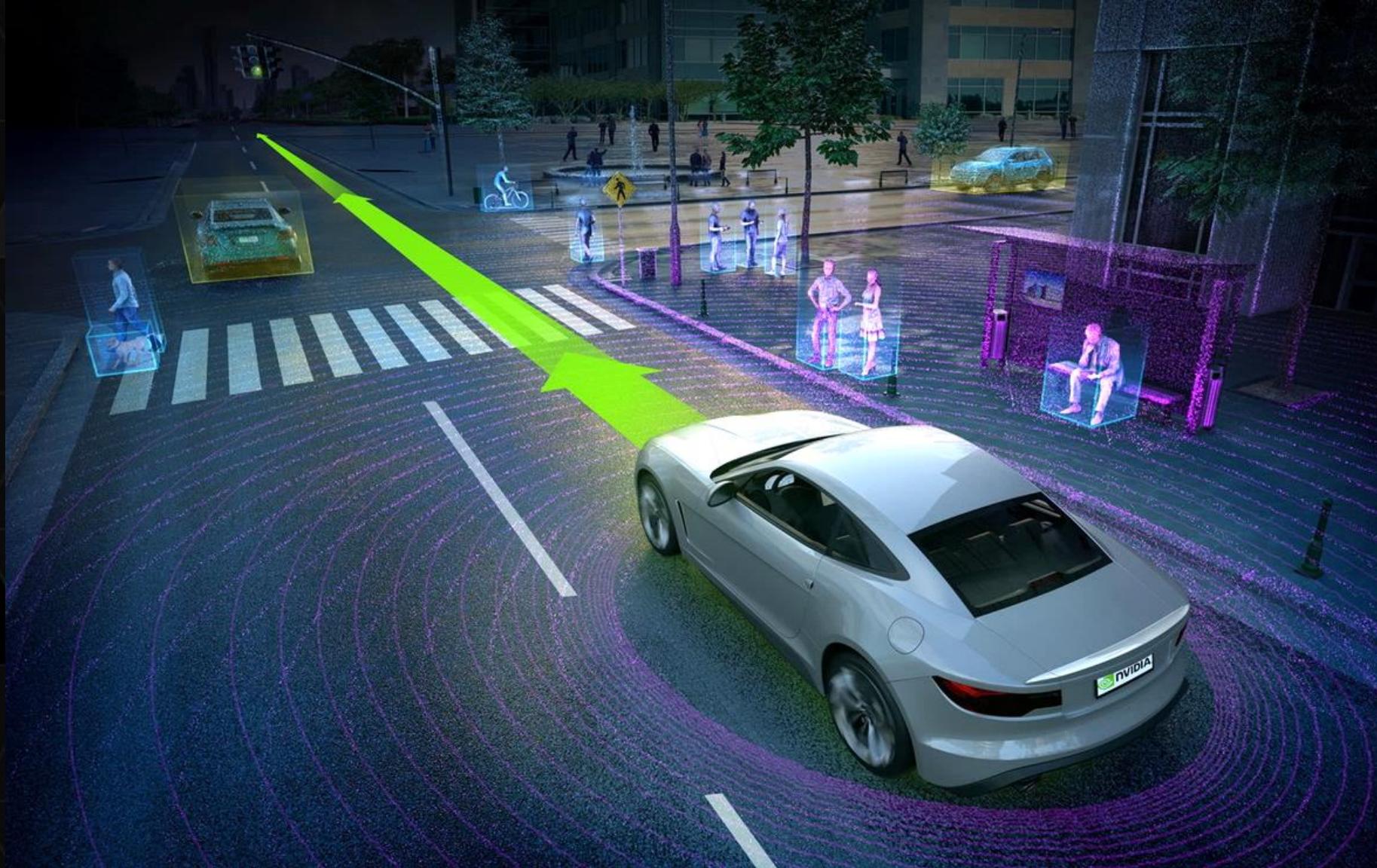


МЕДИЦИНА: ДИАГНОСТИКА

Анализ медицинских изображений и выявление аномалий



АВТОПИЛОТ



ТРУДНОСТИ

Много неструктурированной информации



=

R

0.112	0.986	0.234	0.432	...
0.765	0.128	0.863	0.521	...
1.000	0.985	0.761	0.698	...
0.455	0.783	0.224	0.395	...
0.021	0.500	0.311	0.123	...
1.000	1.000	0.867	0.051	...
1.000	0.945	0.998	0.893	...
0.990	0.941	1.000	0.876	...
0.902	0.867	0.834	0.798	...
.

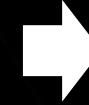
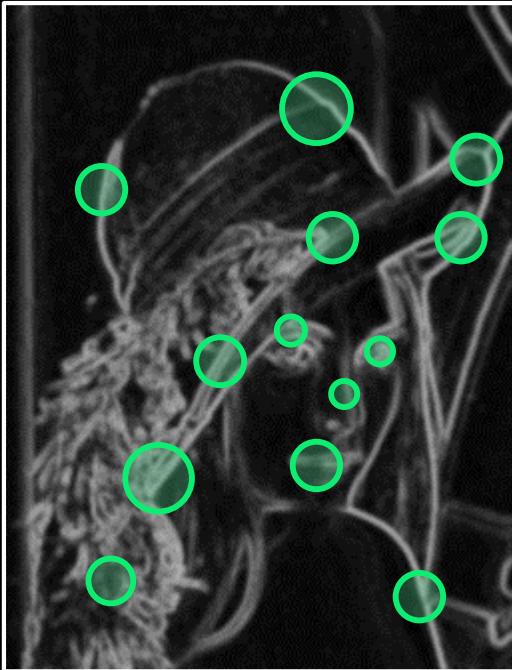
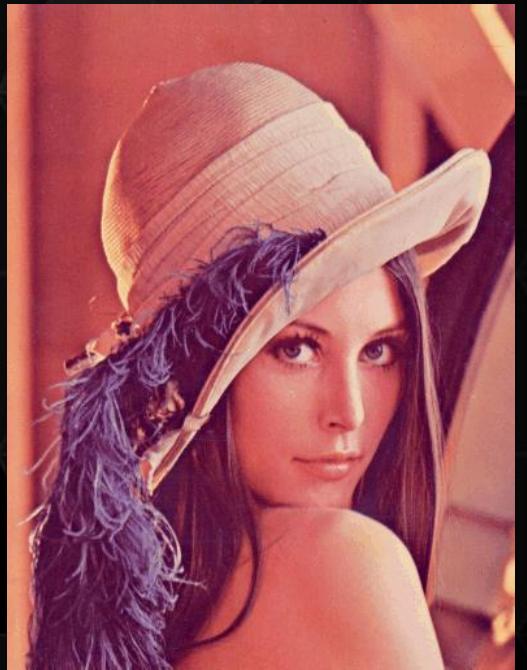
G

0.342	0.647	0.515	0.816	...
0.111	0.300	0.205	0.526	...
0.523	0.428	0.712	0.929	...
0.214	0.604	0.918	0.344	...
0.100	0.121	0.113	0.126	...
.

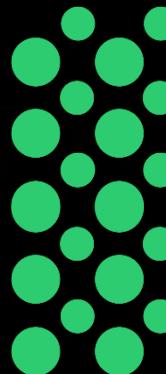
B

0.689	0.706	0.118	0.884	...
0.535	0.532	0.653	0.925	...
0.314	0.265	0.159	0.101	...
0.553	0.633	0.528	0.493	...
0.441	0.465	0.512	0.512	...
.

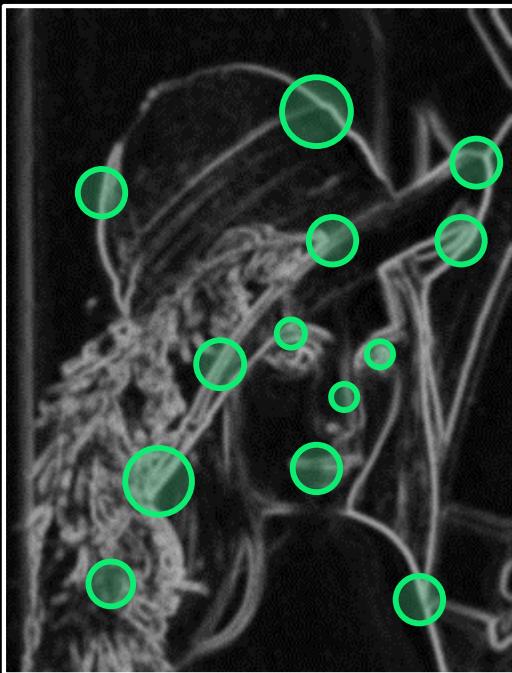
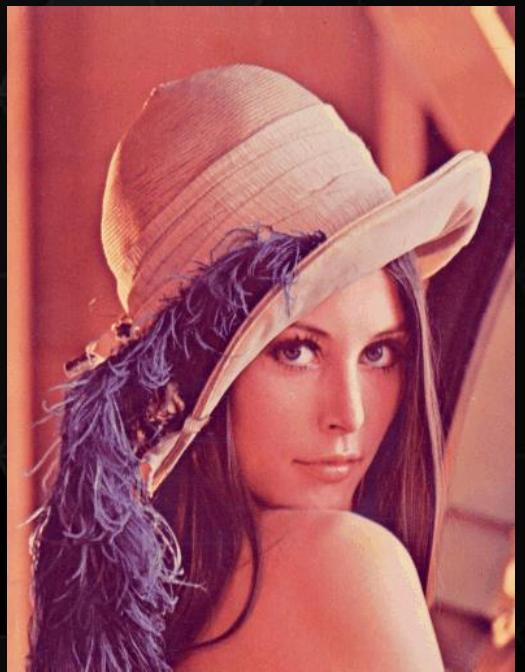
ПРИЗНАКИ В МАШИННОМ ЗРЕНИИ



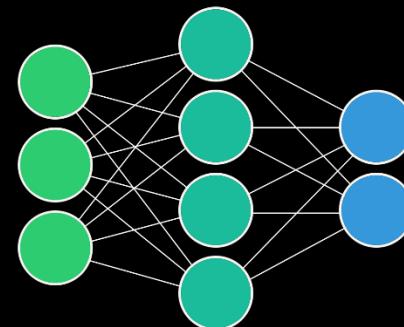
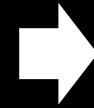
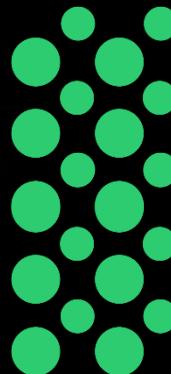
Признаки



ПРИЗНАКИ В МАШИННОМ ЗРЕНИИ



Признаки



“Lenna”

СВЁРТКА

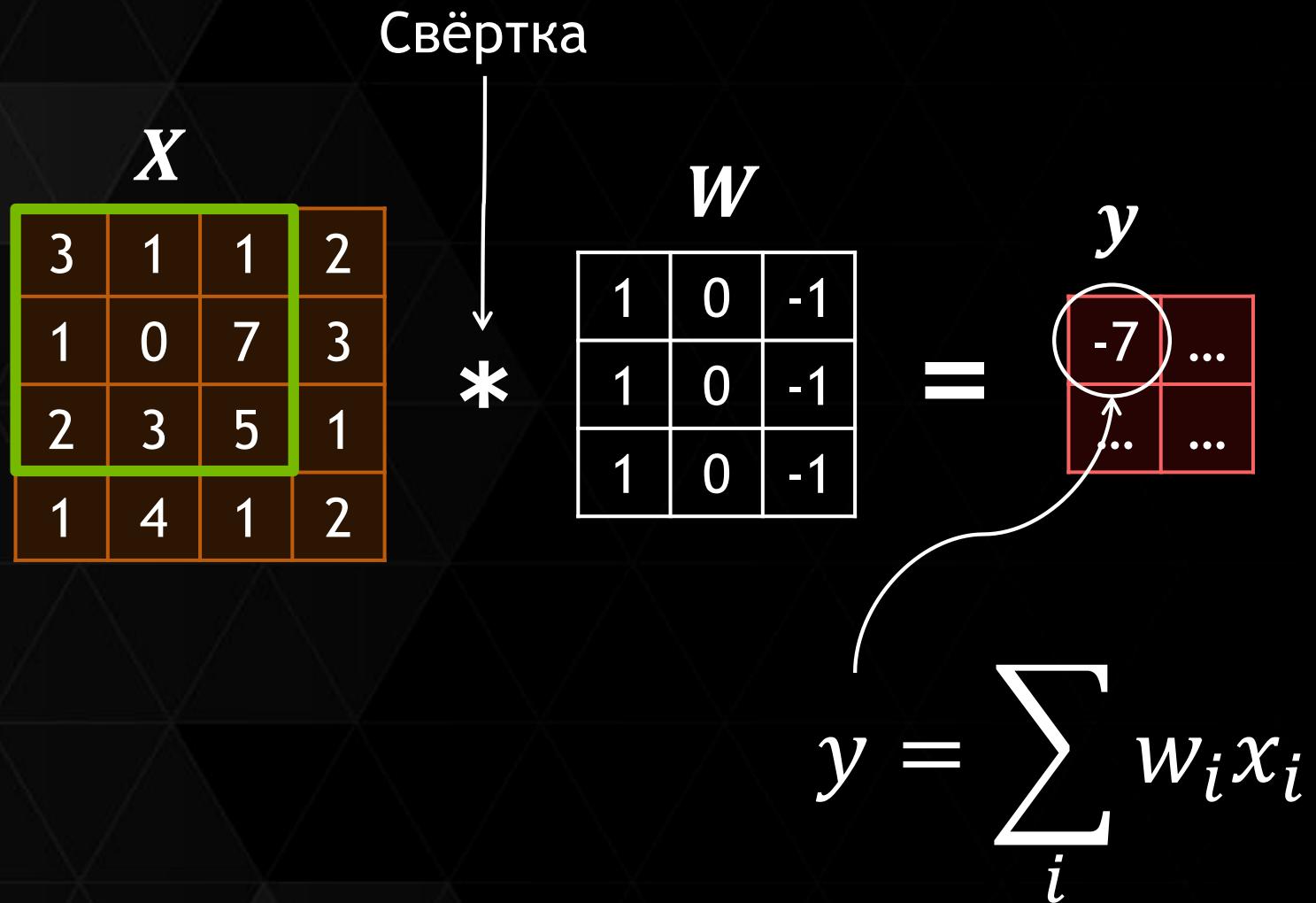
X

3	1	1	2
1	0	7	3
2	3	5	1
1	4	1	2

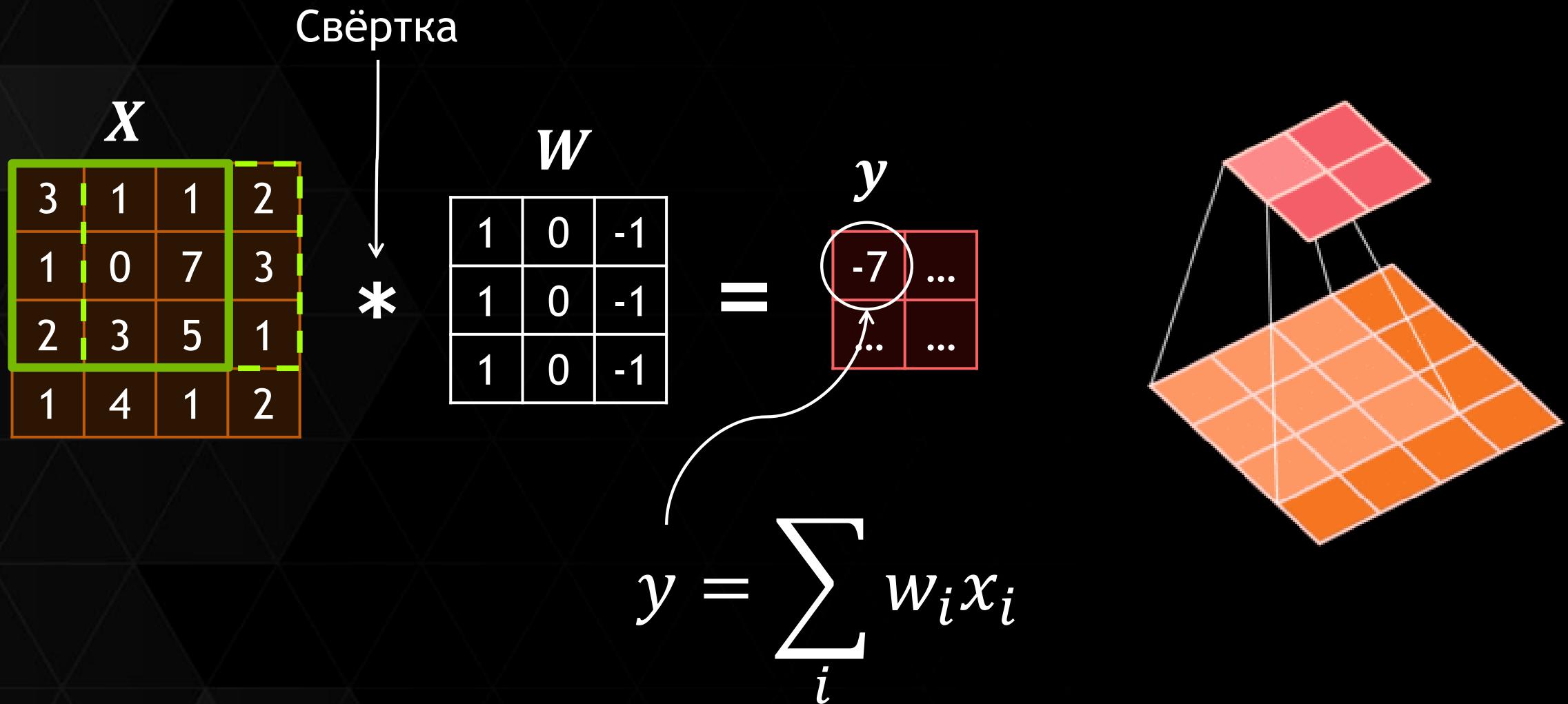
W

1	0	-1
1	0	-1
1	0	-1

СВЁРТКА



СВЁРТКА

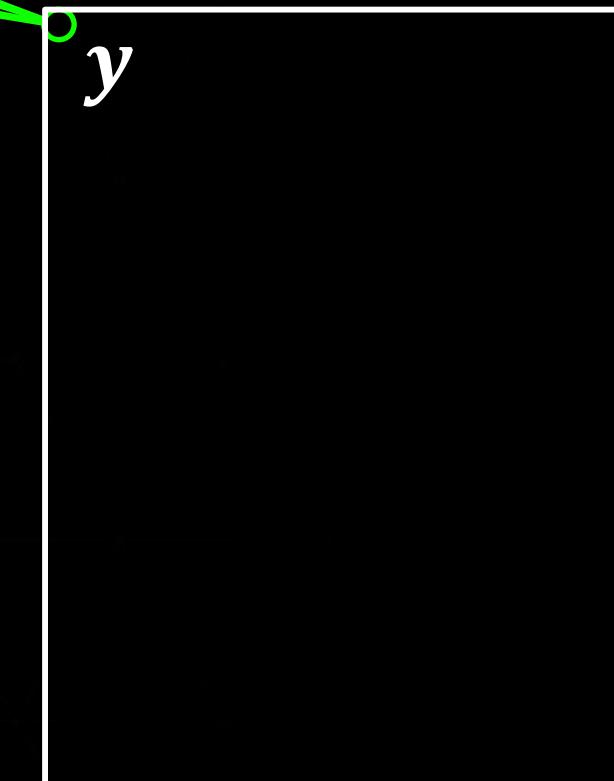


СВЁРТКА

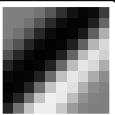
$$W \quad *$$



$$y = \sum_i w_i x_i$$



СВЁРТКА

$$W$$


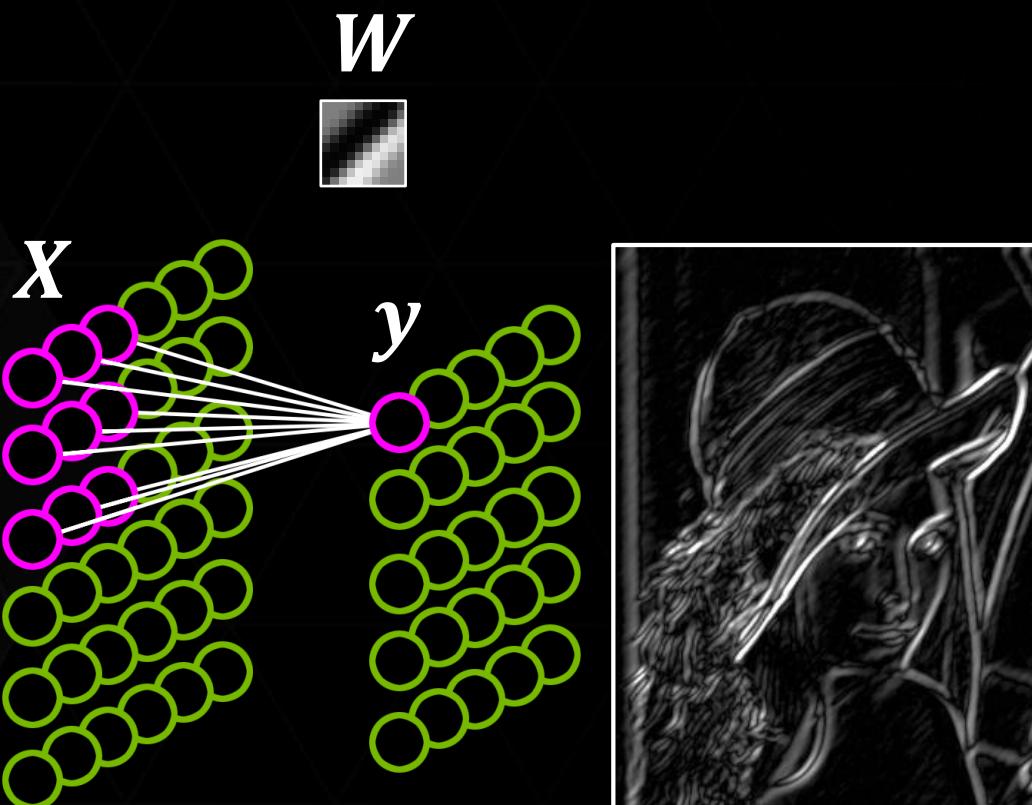
*



$$y = \sum_i w_i x_i$$

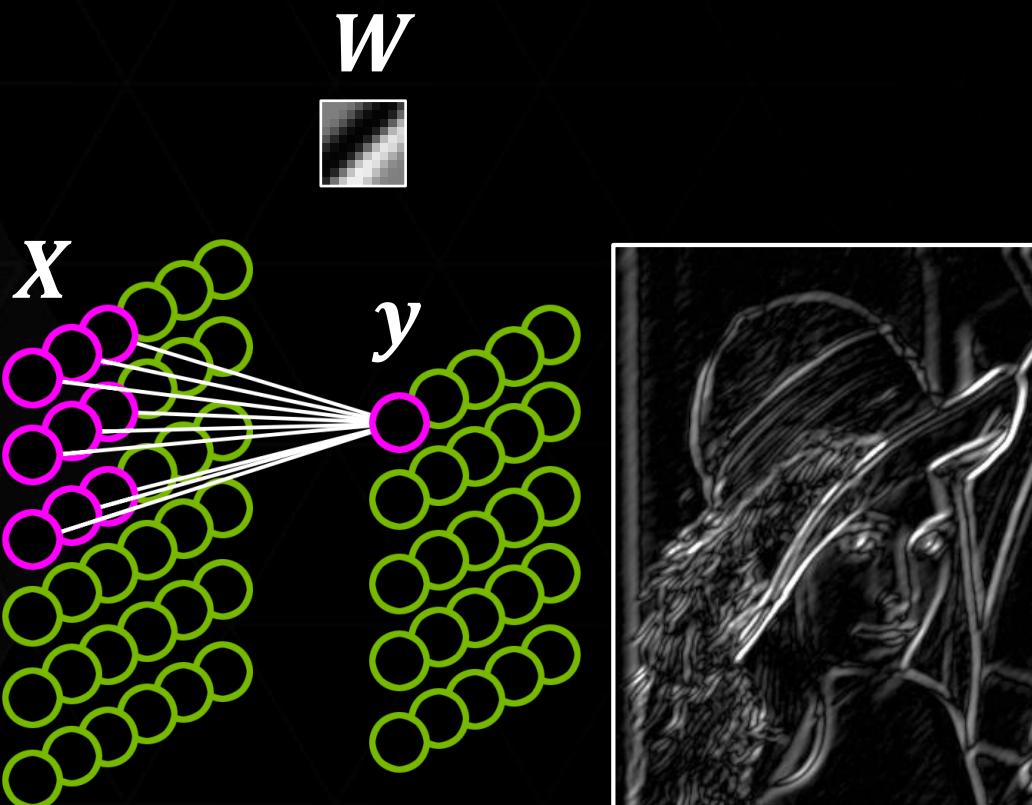


СВЁРТОЧНЫЙ СЛОЙ НЕЙРОННОЙ СЕТИ



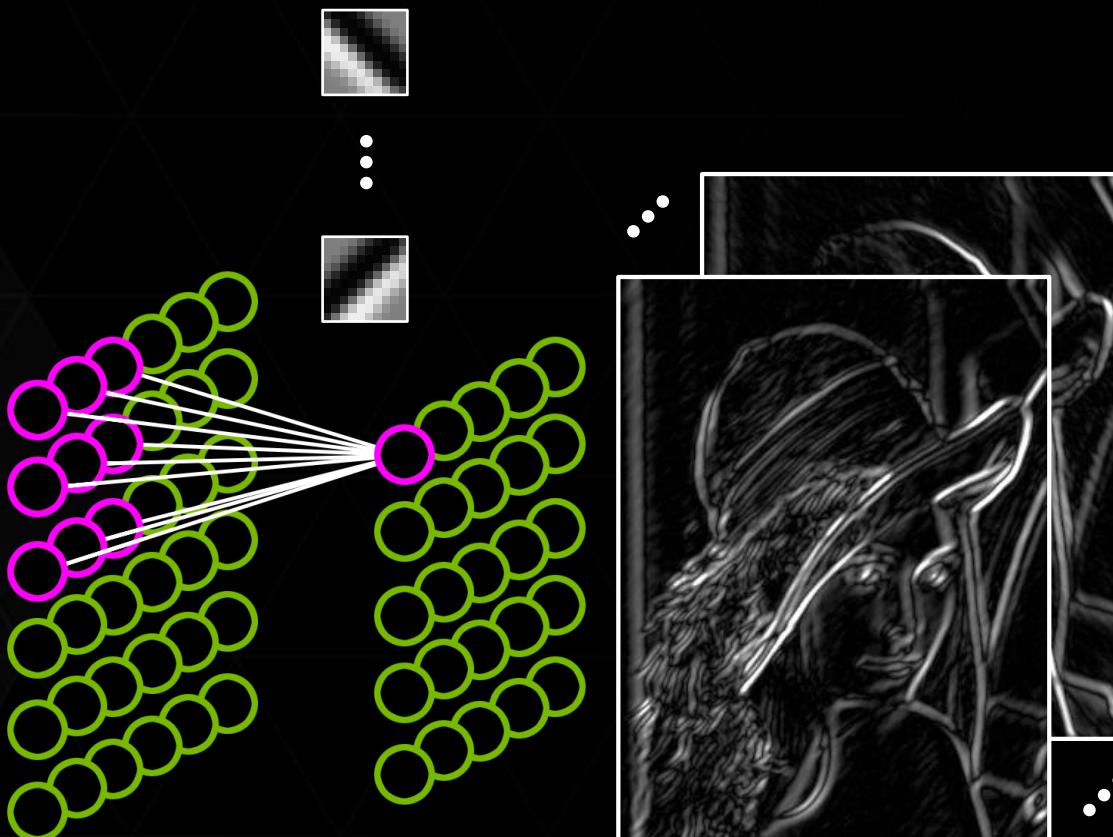
$$y = \sum_i w_i x_i$$

СВЁРТОЧНЫЙ СЛОЙ НЕЙРОННОЙ СЕТИ



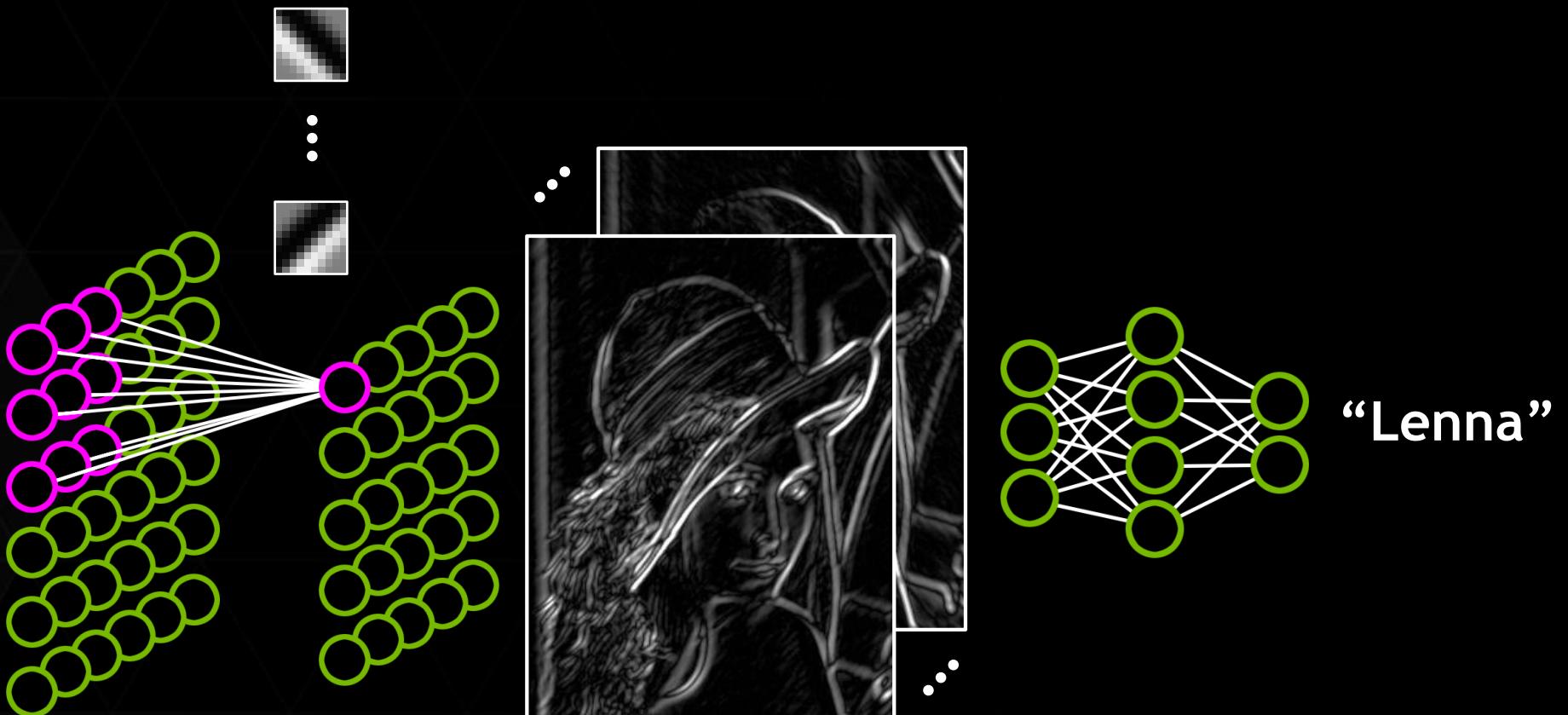
$$y = f \left(\sum_i w_i x_i + b \right)$$

СВЁРТОЧНЫЙ СЛОЙ НЕЙРОННОЙ СЕТИ



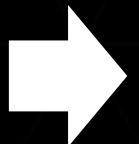
Обучаемые
признаки

СВЁРТОЧНАЯ НЕЙРОННАЯ СЕТЬ



Обучаемые
признаки

ПРОБЛЕМЫ НЕ ЗАКОНЧИЛИСЬ



“шесть”



Легко

ПРОБЛЕМЫ НЕ ЗАКОНЧИЛИСЬ



“шесть”



Легко



?
=



Сложно

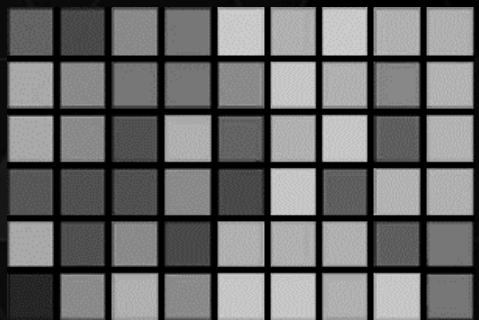
A close-up shot from the movie Inception. Two men in dark suits are seated at a table, looking down at it with serious expressions. The lighting is dramatic, casting deep shadows on their faces. The man on the left has light-colored hair and is wearing a white shirt and a dark tie. The man on the right has dark hair and is also wearing a white shirt and a dark tie. The background is blurred, showing what appears to be a train car interior.

WE NEED TO GO

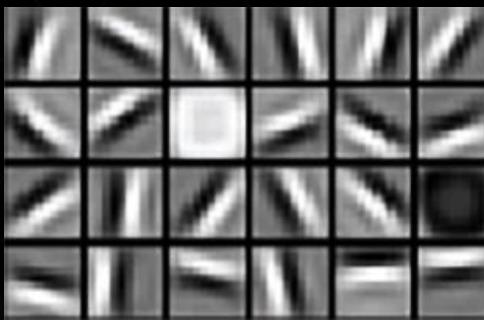
DEEPER

ИЕРАРХИЯ ПРИЗНАКОВ

Пиксели



Границы



Фрагменты

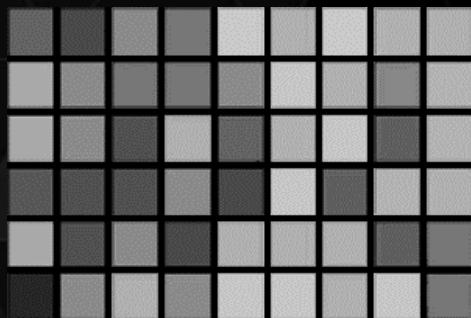


Объекты

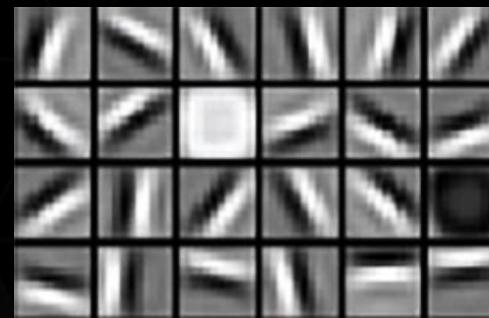


ИЕРАРХИЯ ПРИЗНАКОВ

Пиксели



Границы



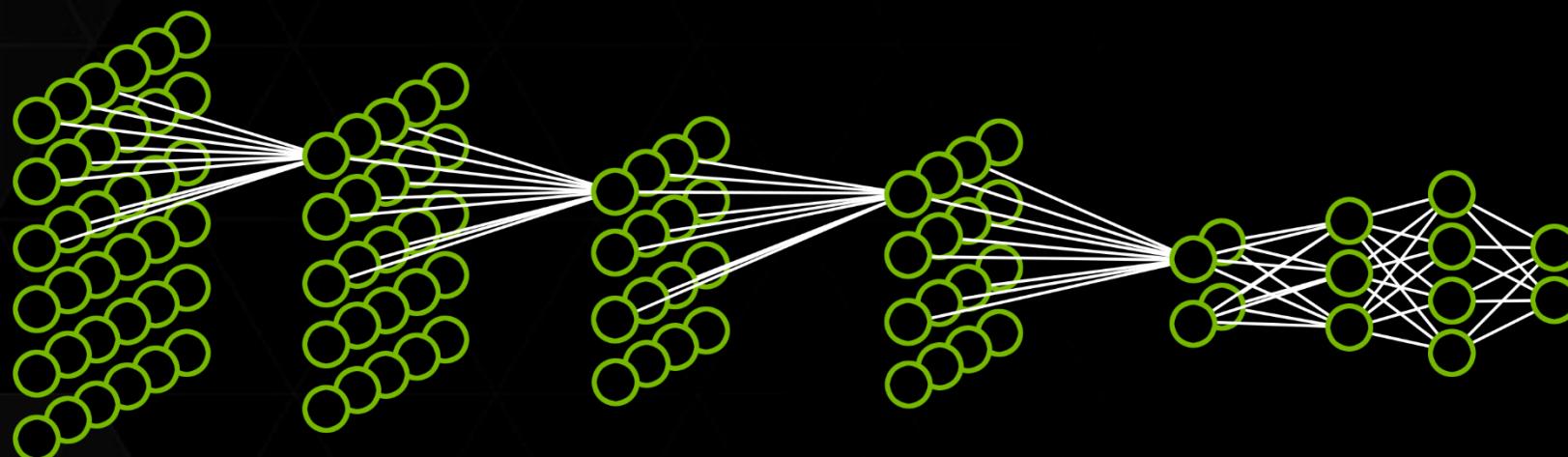
Фрагменты



Объекты



ГЛУБОКИЕ НЕЙРОННЫЕ СЕТИ



“Котик”

DEEP LEARNING

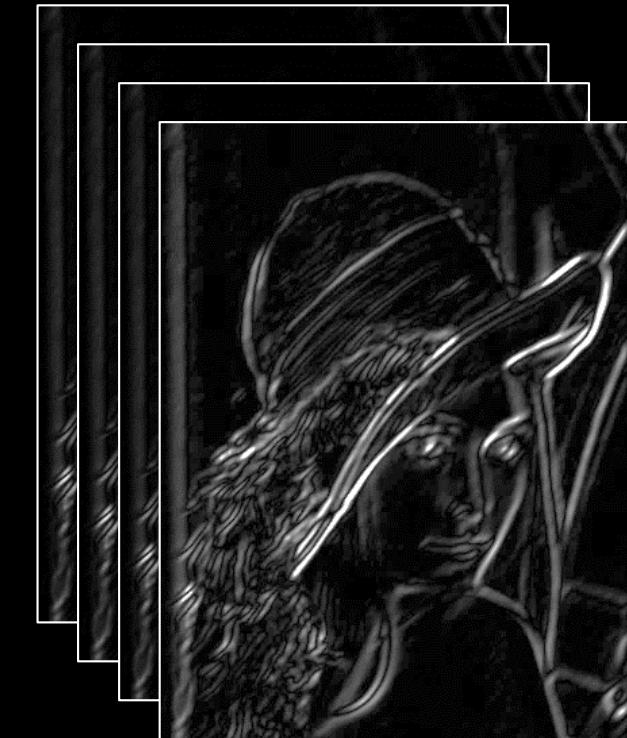
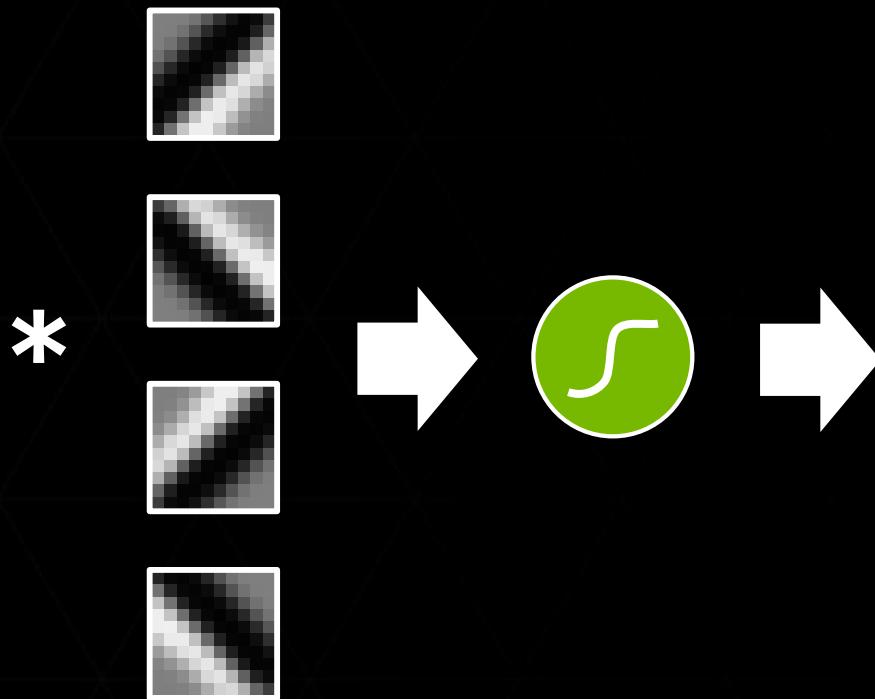
Обучение иерархии признаковых представлений

DEEP LEARNING

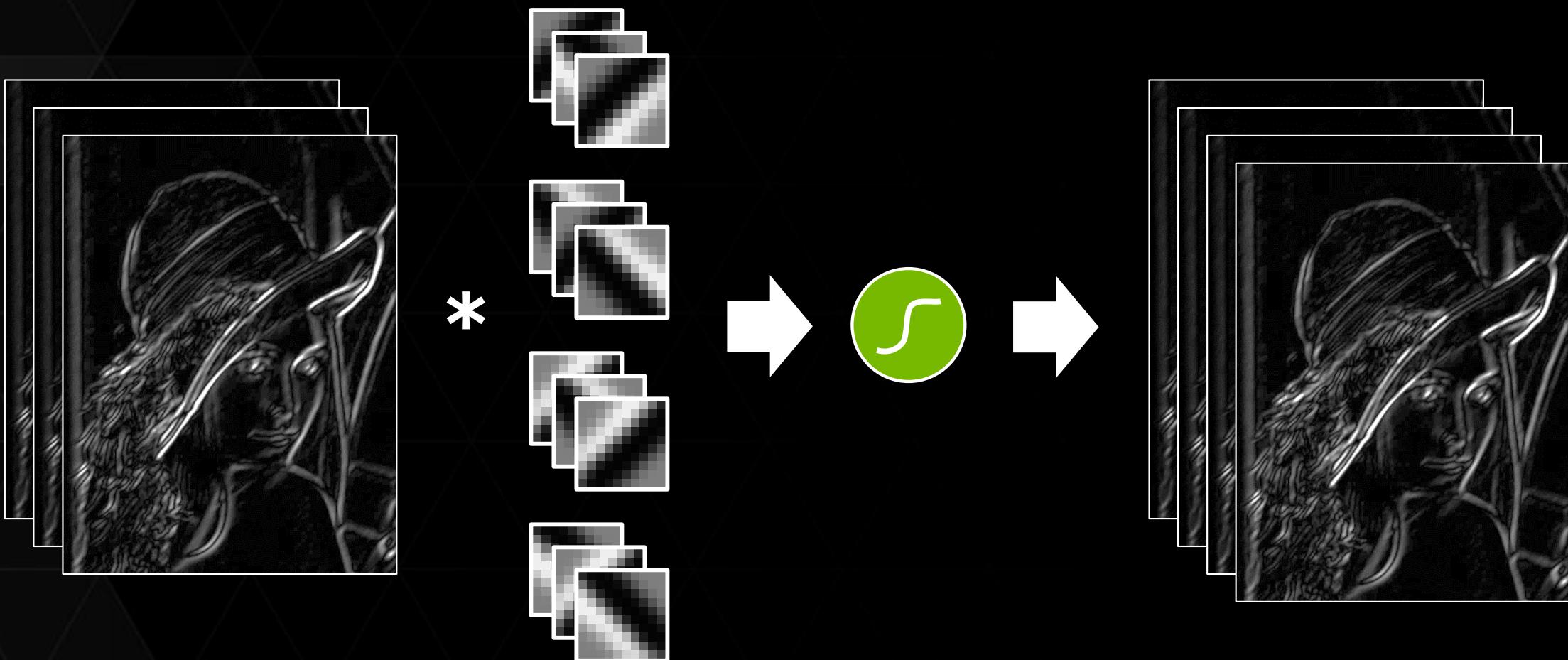
Обучение иерархии признаковых представлений

- ✓ Признаки
- ✓ Обучаемые признаки
- ✓ Иерархия (много уровней)

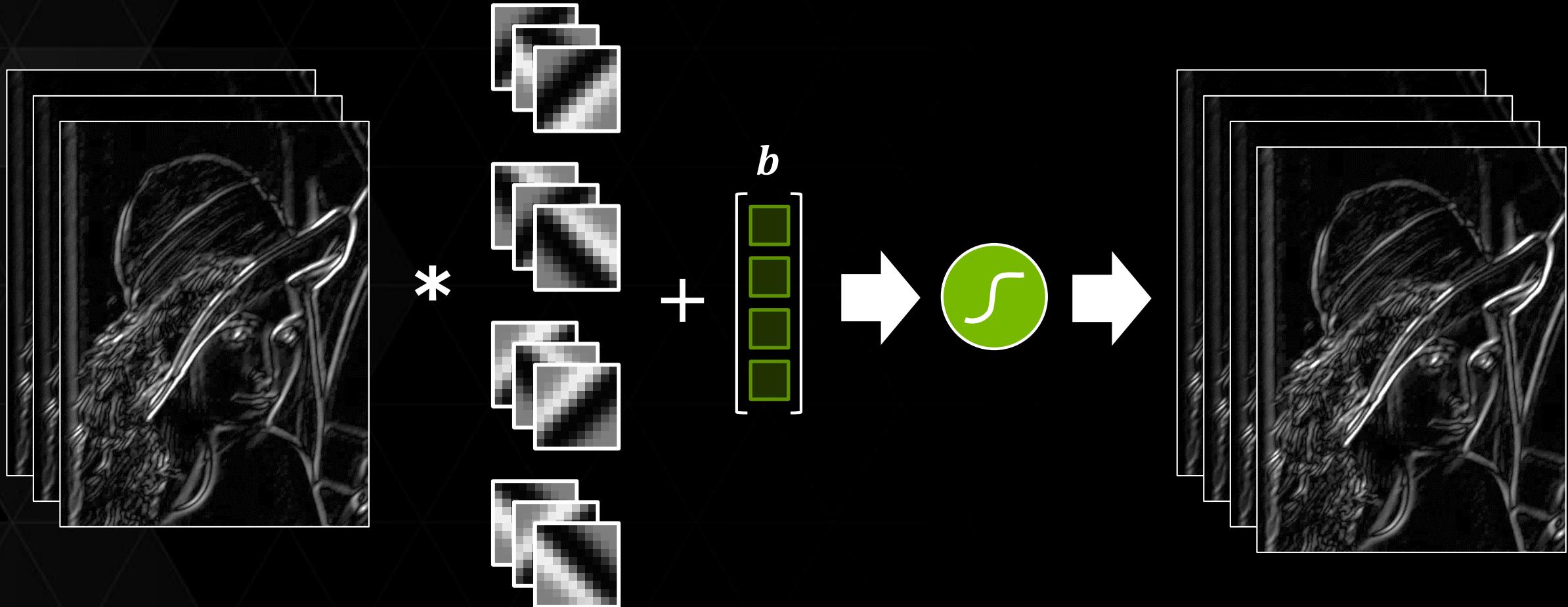
СВЁРТОЧНЫЙ СЛОЙ (CONV)



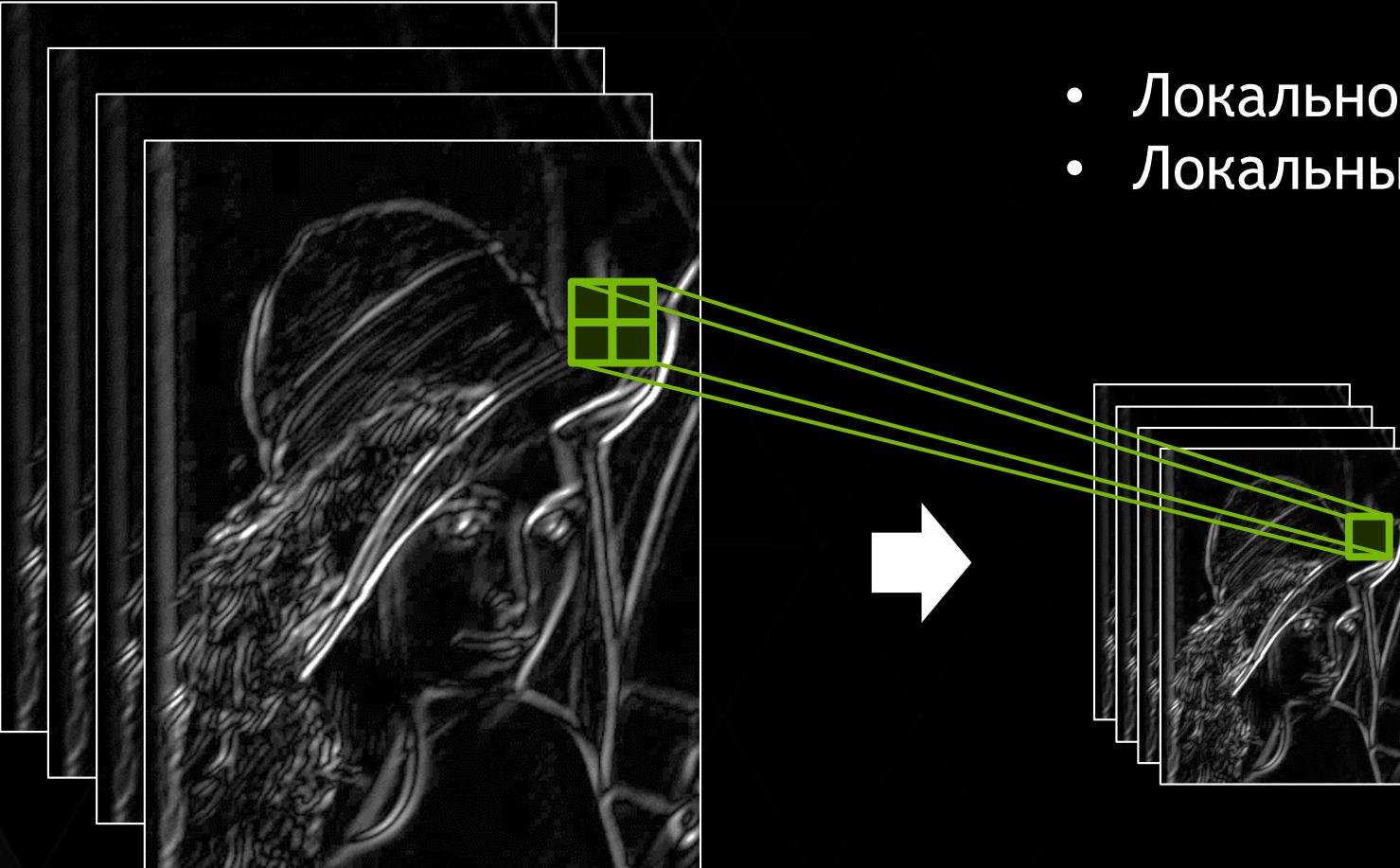
СВЁРТОЧНЫЙ СЛОЙ (CONV)



СВЁРТОЧНЫЙ СЛОЙ (CONV)



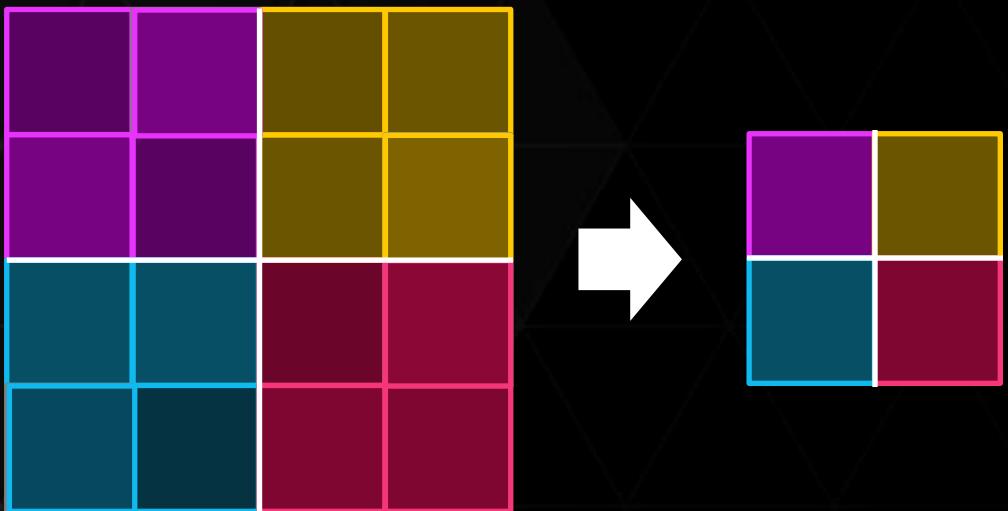
ПОНИЖЕНИЕ РАЗМЕРНОСТИ (POOLING)



- Локальное усреднение
- Локальный максимум

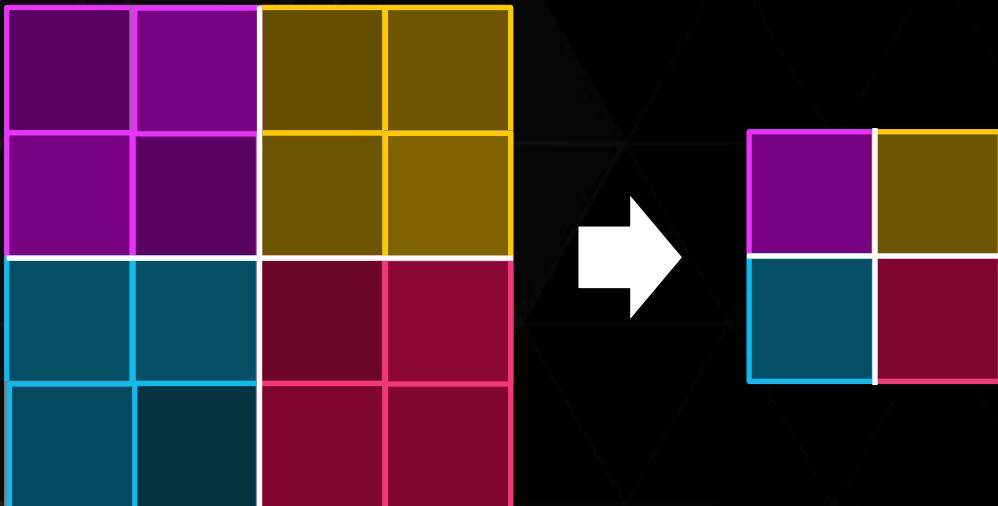
ПОНИЖЕНИЕ РАЗМЕРНОСТИ (POOLING)

Average Pooling

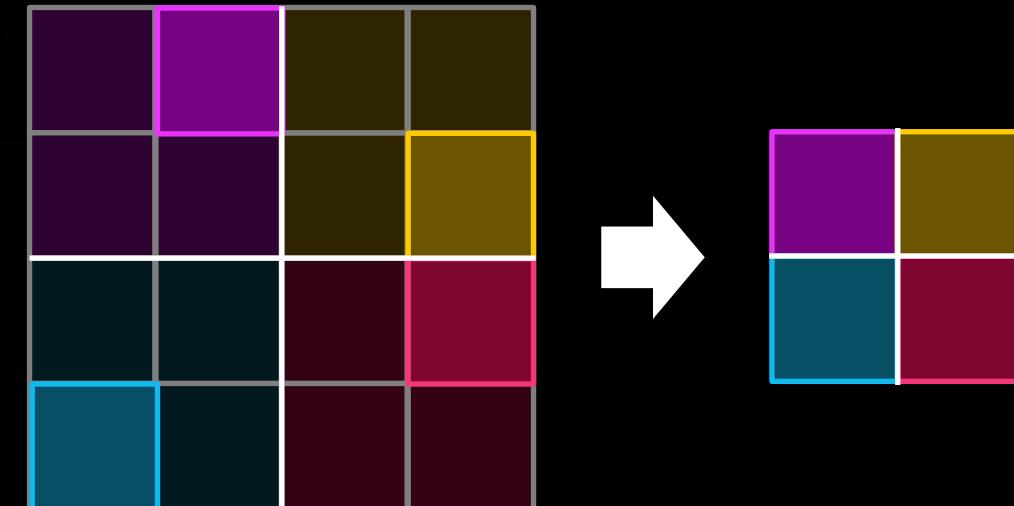


ПОНИЖЕНИЕ РАЗМЕРНОСТИ (POOLING)

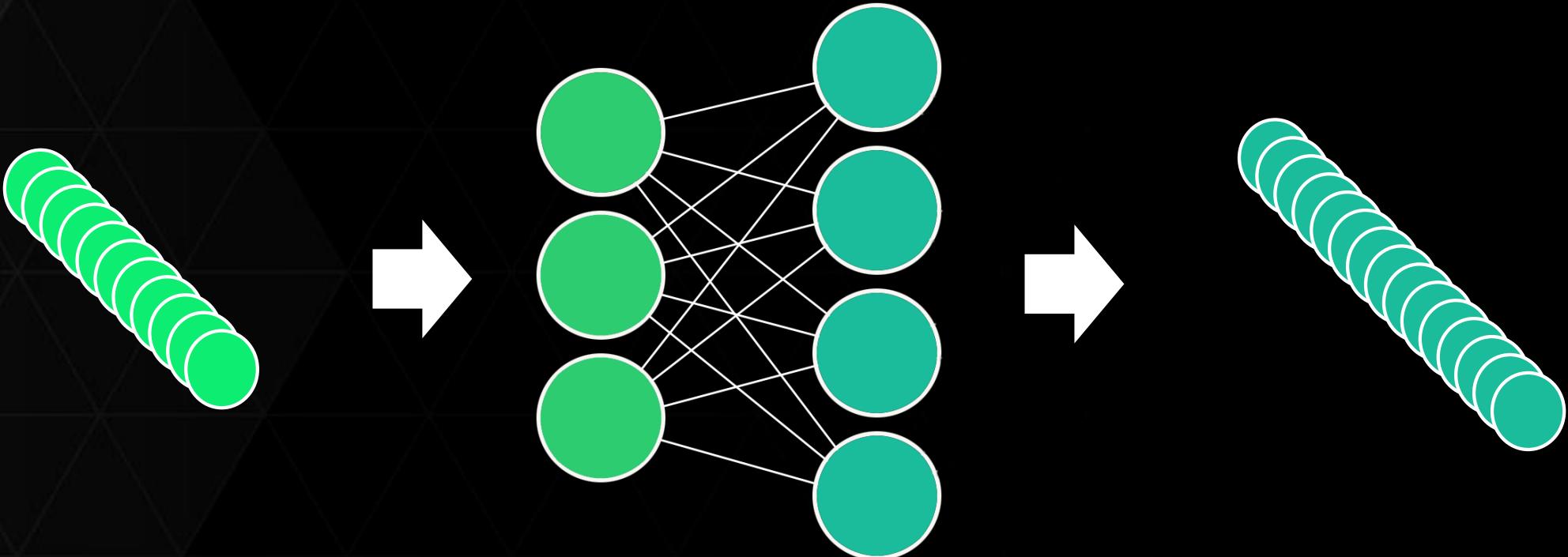
Average Pooling



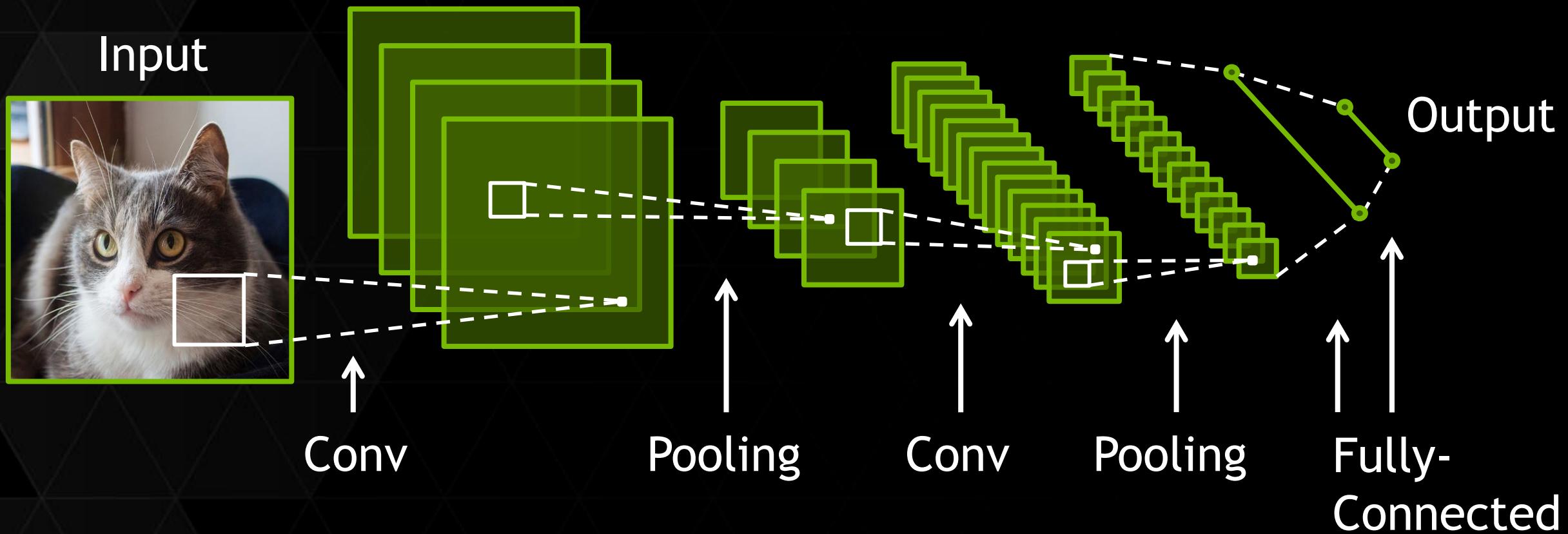
Max Pooling



ПОЛНОСВЯЗНЫЙ СЛОЙ (FULLY-CONNECTED)

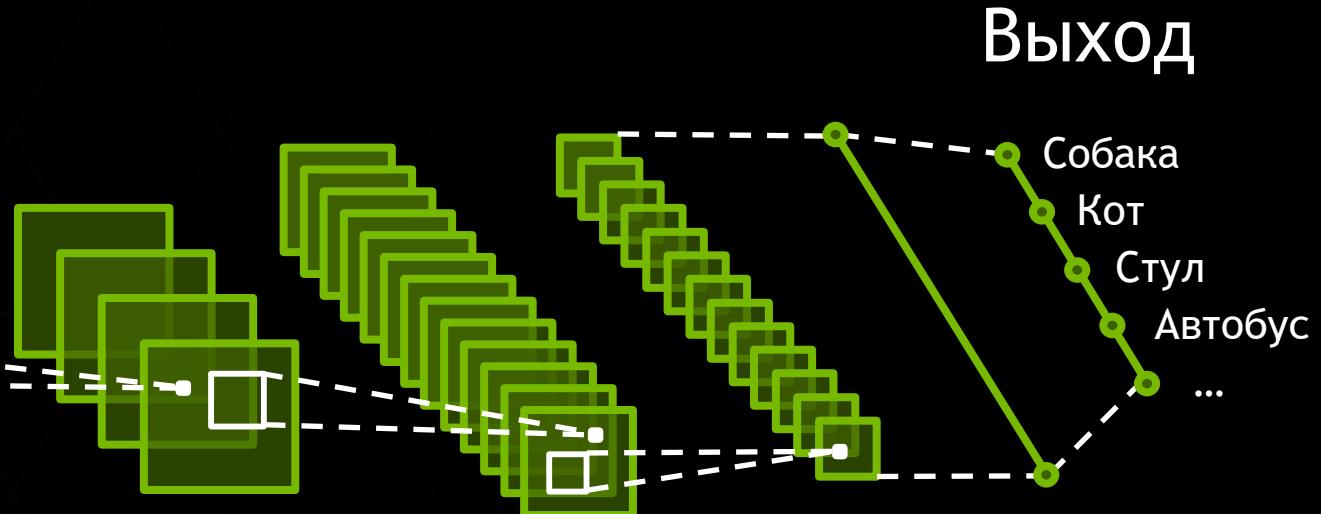
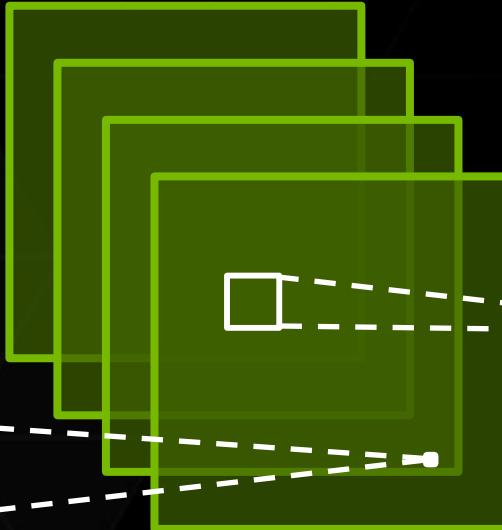
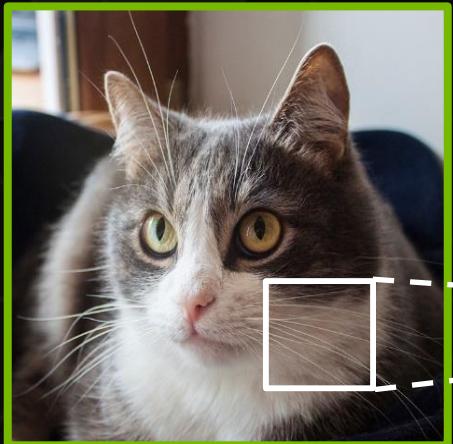


ПРИМЕР СВЁРТОЧНОЙ НЕЙРОННОЙ СЕТИ

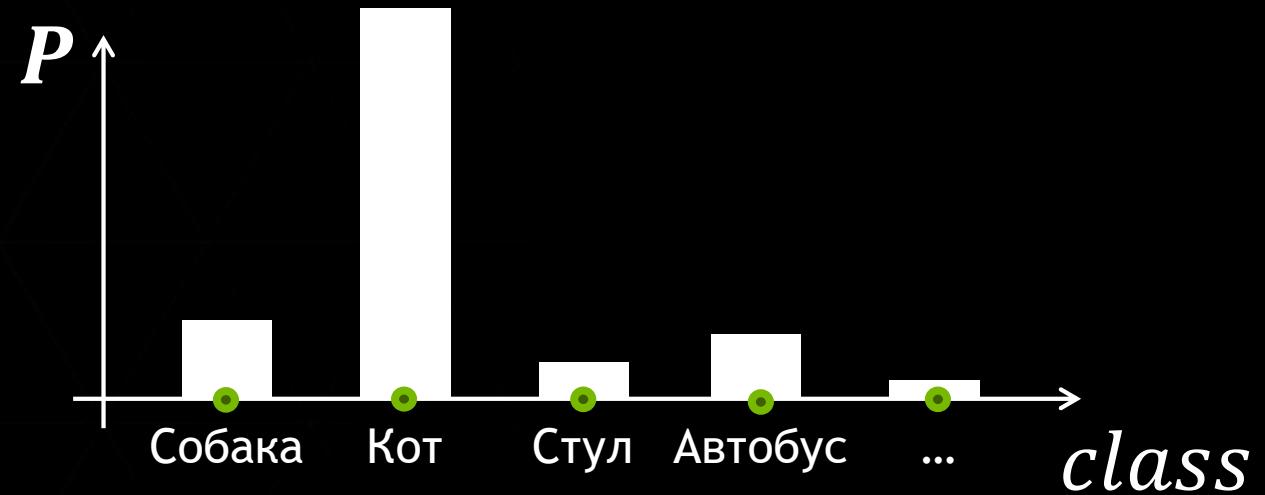


КЛАССИФИКАЦИЯ ИЗОБРАЖЕНИЙ

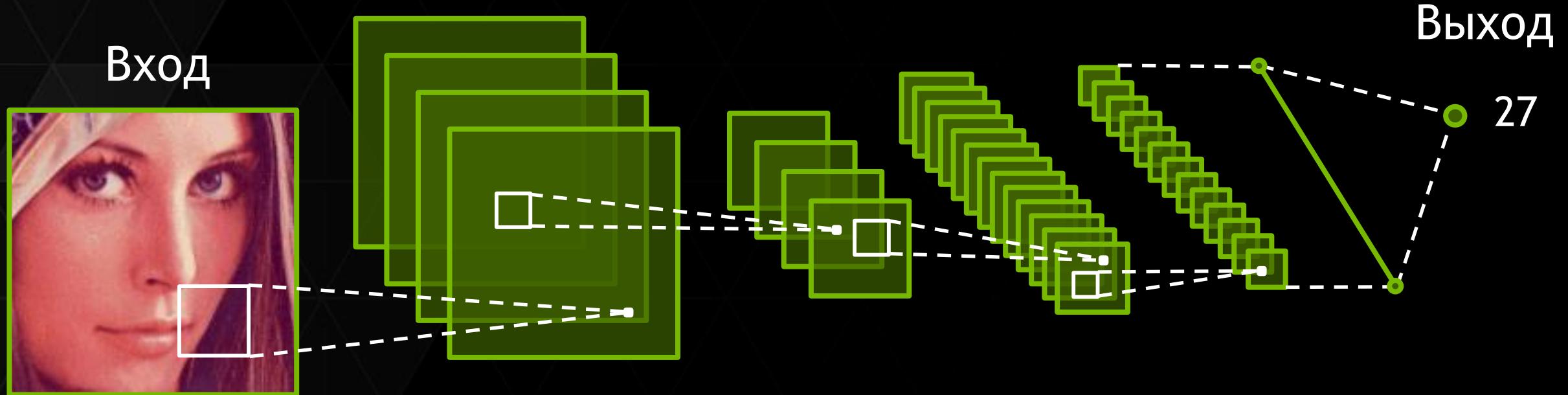
Вход



Выход сети -
распределение вероятностей
принадлежности к классу



РЕГРЕССИЯ НА ИЗОБРАЖЕНИЯХ



Выход сети -
непрерывное значение
(или вектор/тензор)

ОБРАБОТКА ПОСЛЕДОВАТЕЛЬНОСТЕЙ

ОБРАБОТКА ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Мама

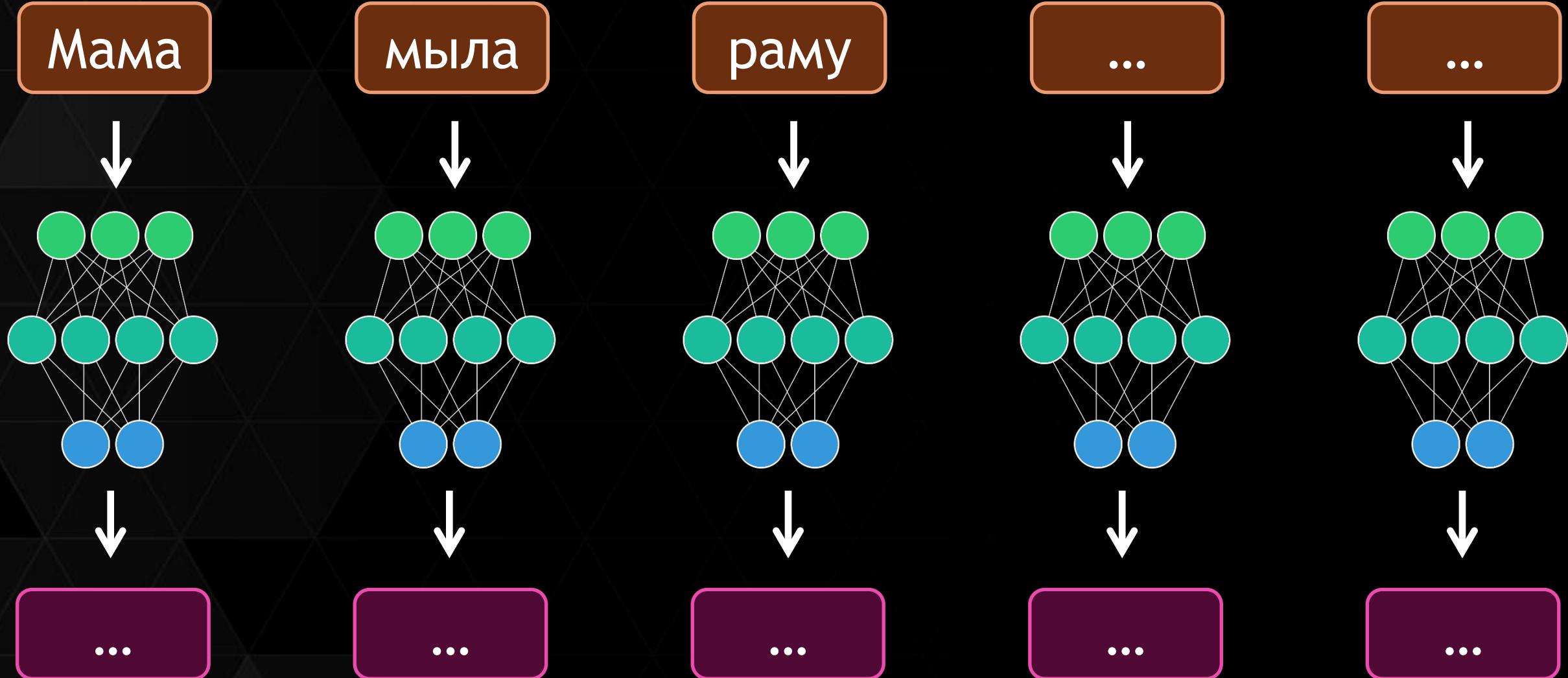
мыла

раму

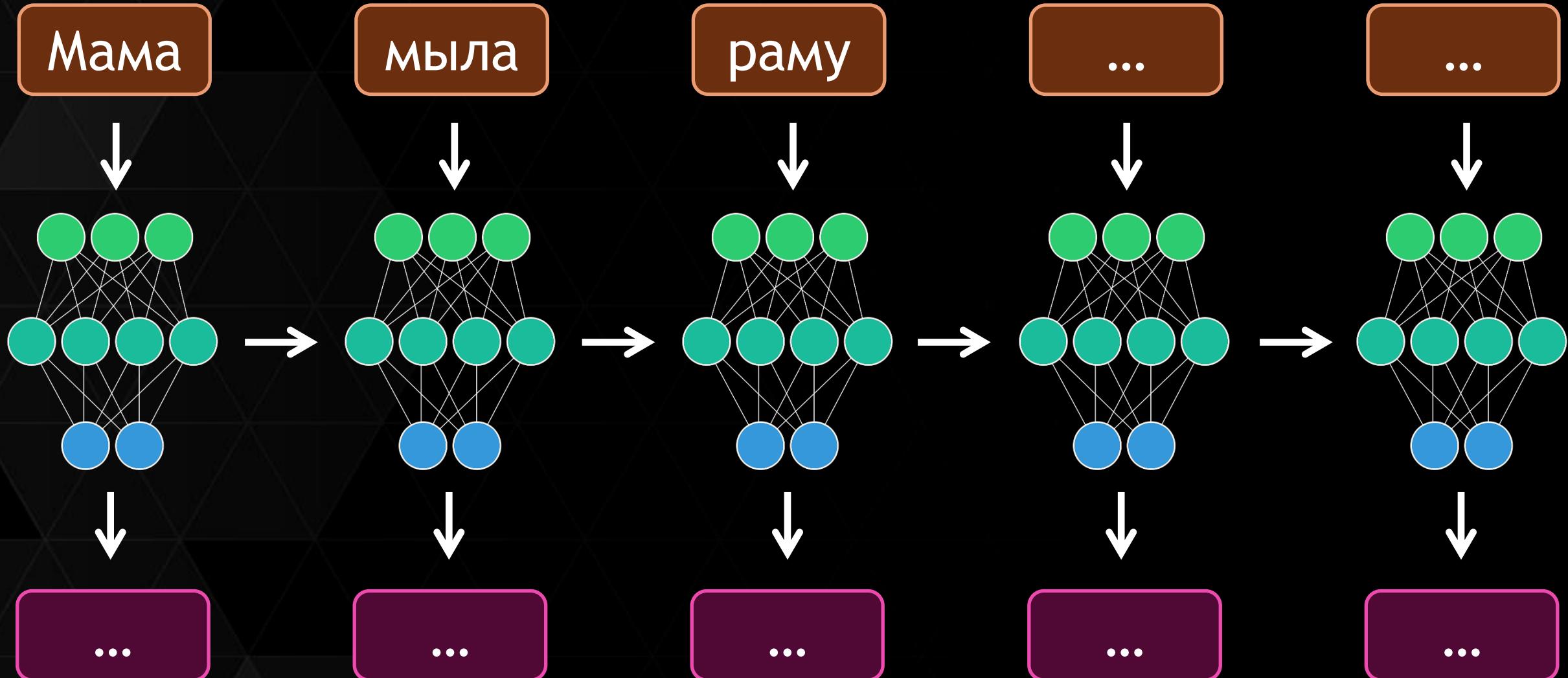
...

...

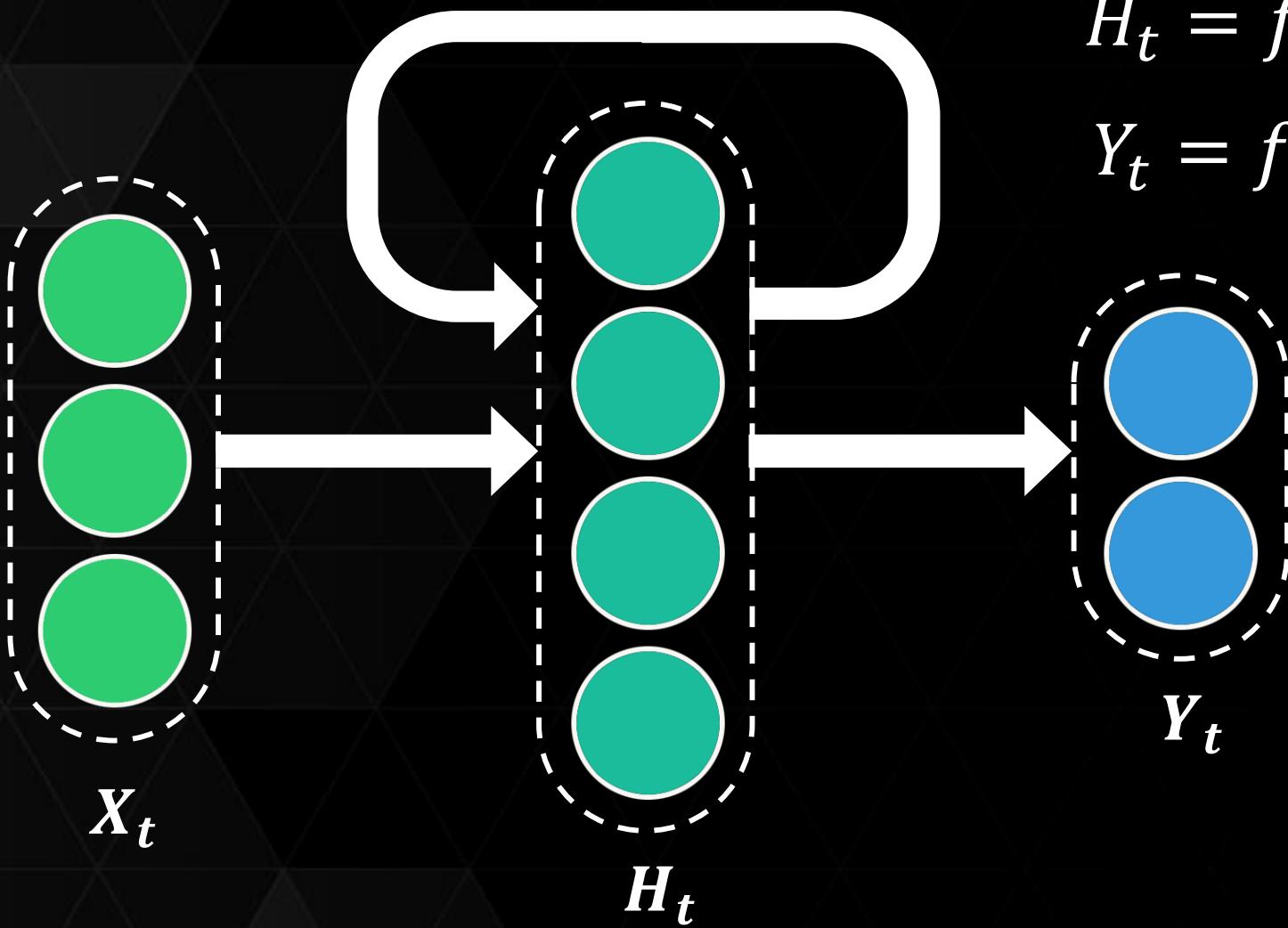
ОБРАБОТКА ПОСЛЕДОВАТЕЛЬНОСТЕЙ



РЕКУРРЕНТНЫЕ НЕЙРОННЫЕ СЕТИ



РЕКУРРЕНТНЫЕ НЕЙРОННЫЕ СЕТИ

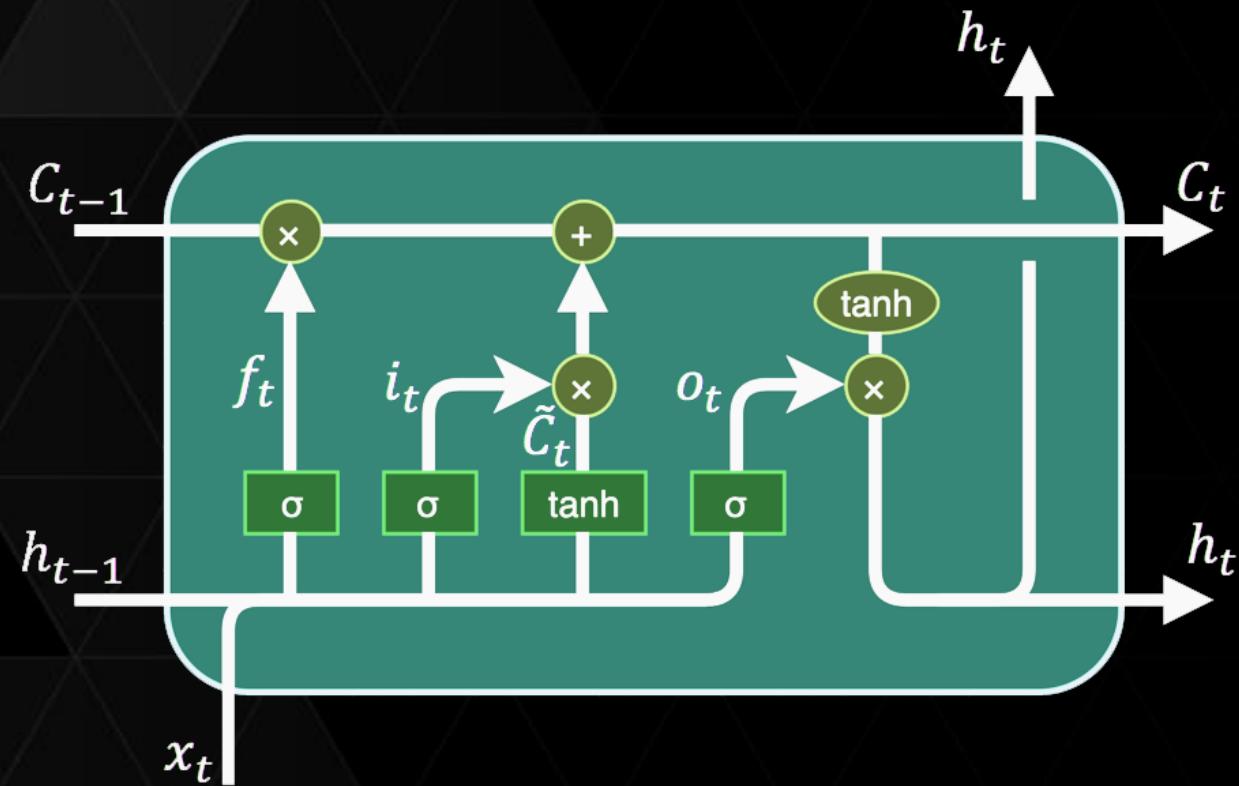


$$H_t = f(W_1 X_t + W_h H_{t-1} + B_1)$$

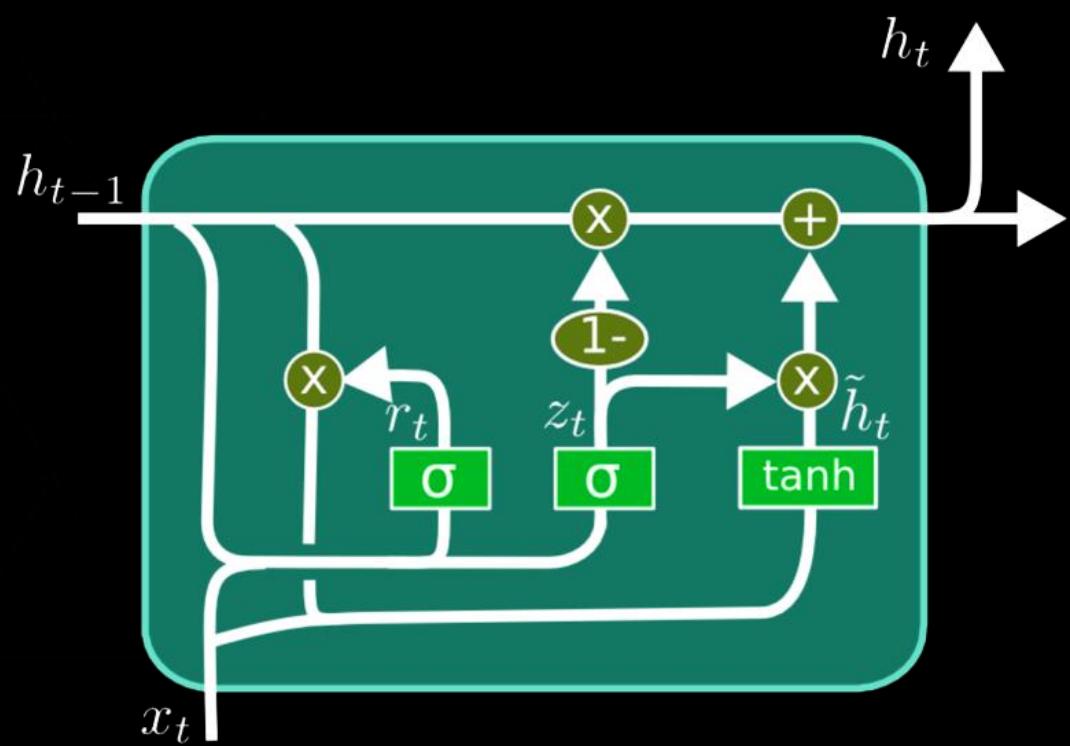
$$Y_t = f(W_2 H_t + B_2)$$

РЕКУРРЕНТНЫЕ НЕЙРОННЫЕ СЕТИ

LSTM

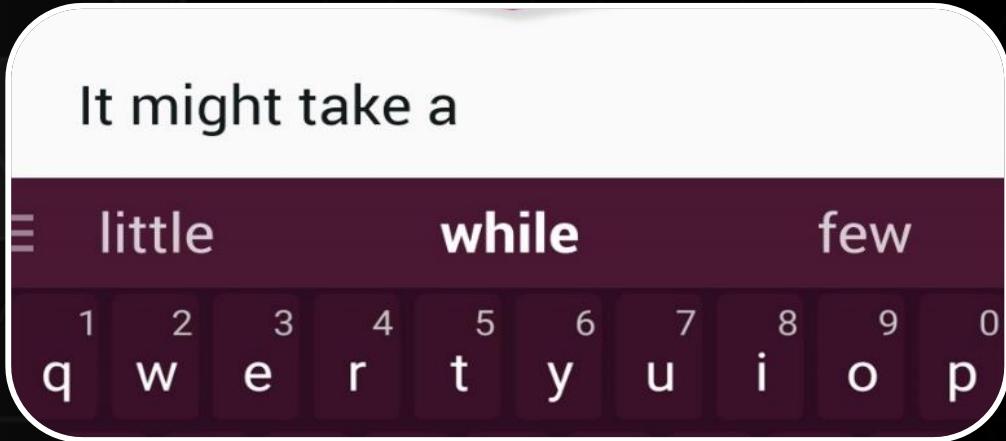


GRU



ОБРАБОТКА ЕСТЕСТВЕННОГО ЯЗЫКА

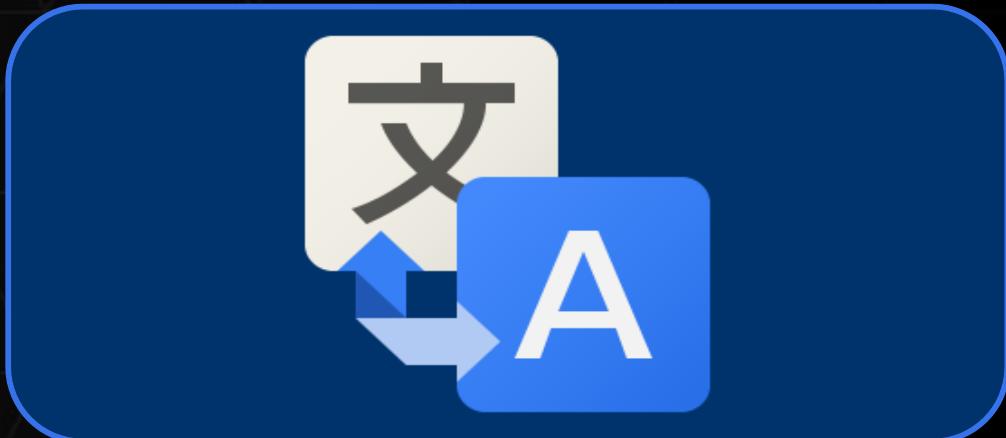
Умная клавиатура



Анализ комментариев



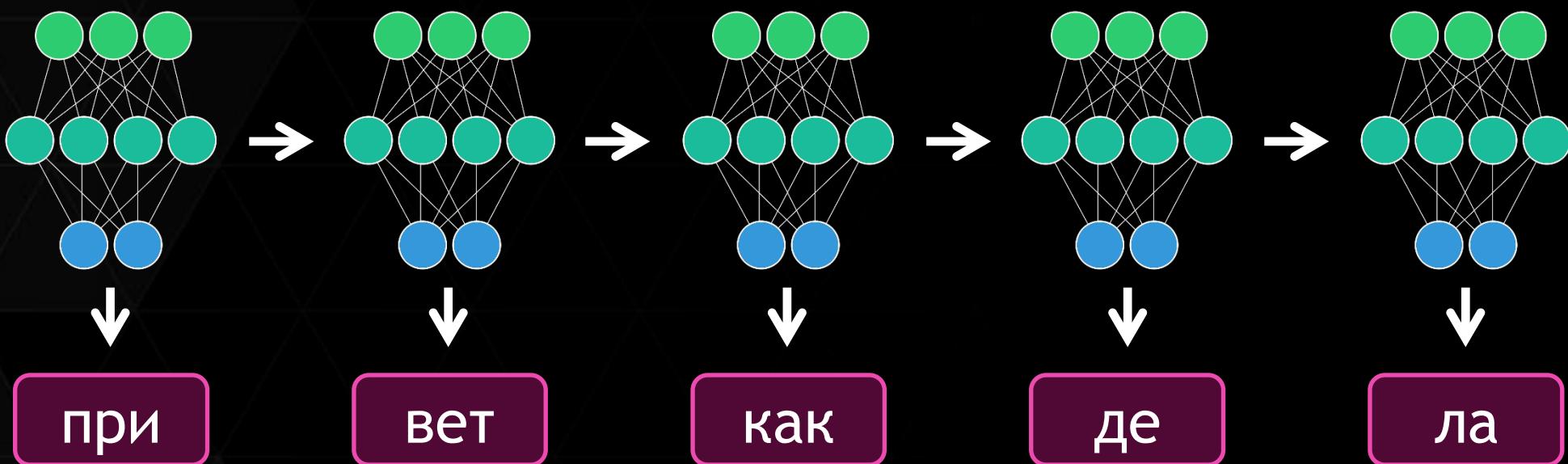
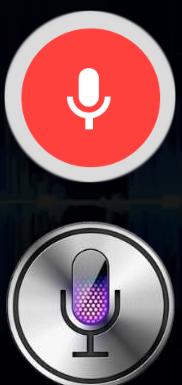
Машинный перевод



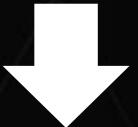
Чат боты



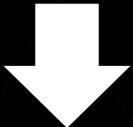
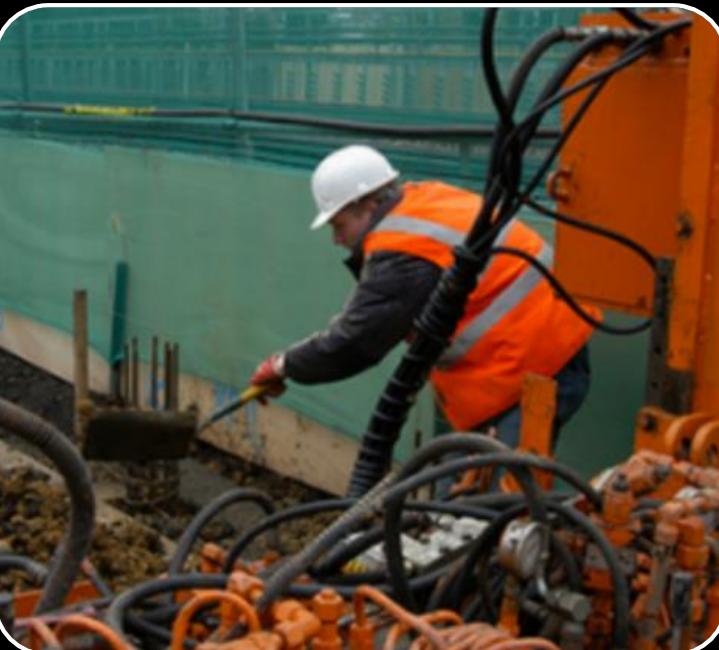
РАСПОЗНАВАНИЕ РЕЧИ



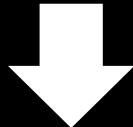
ИЗОБРАЖЕНИЕ → ТЕКСТ



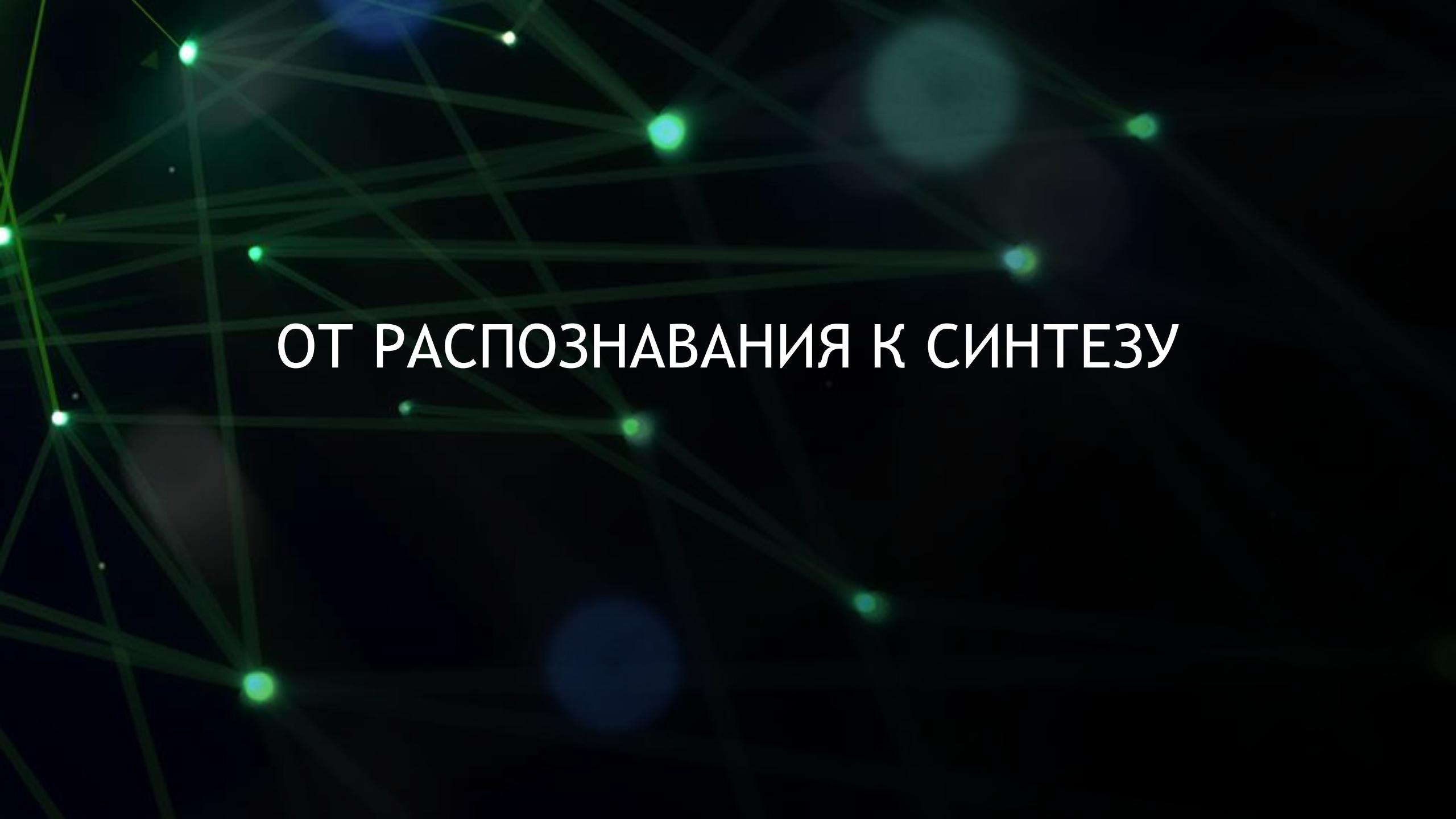
Man in black shirt is
playing guitar



Construction worker
in orange safety vest
is working on road

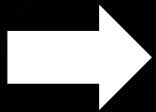


Two young girls are
playing with lego
toy



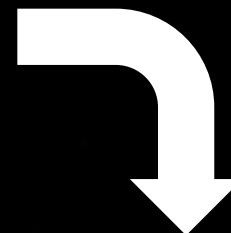
ОТ РАСПОЗНАВАНИЯ К СИНТЕЗУ

ВОСПРИЯТИЕ И ТВОРЧЕСТВО



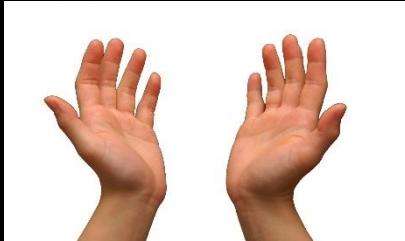
Интерфейс

Восприятие

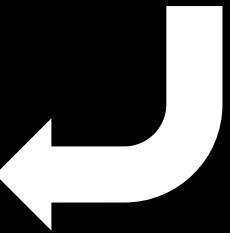
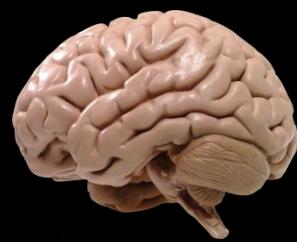


Концепт

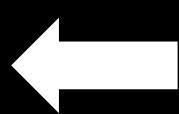
ВОСПРИЯТИЕ И ТВОРЧЕСТВО



Интерфейс



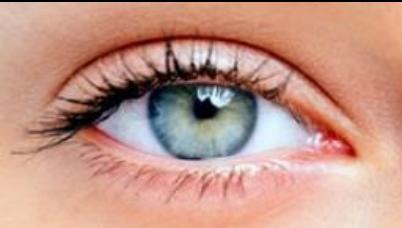
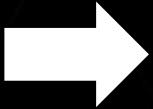
Творчество



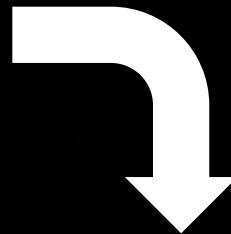
Концепт



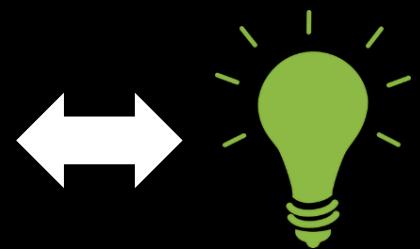
ВОСПРИЯТИЕ И ТВОРЧЕСТВО



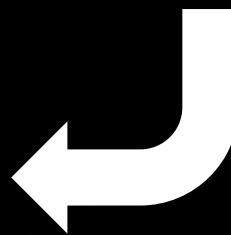
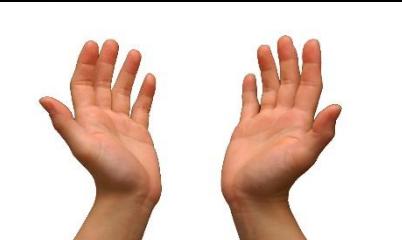
Восприятие



Концепт

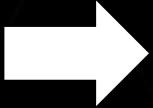


Интерфейс

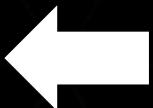


Творчество

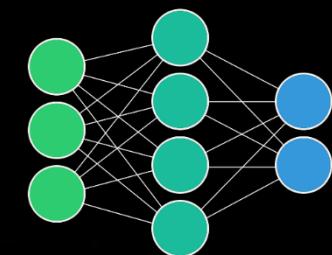
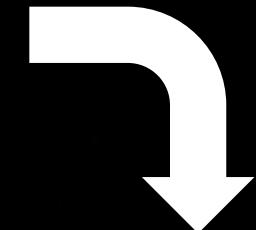
РАСПОЗНАВАНИЕ И СИНТЕЗ



Распознавание



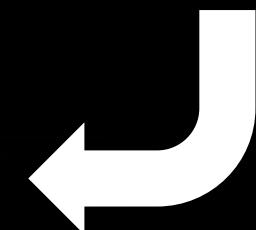
Интерфейс



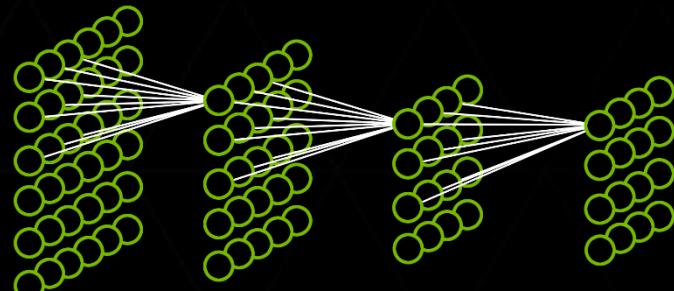
Концепт



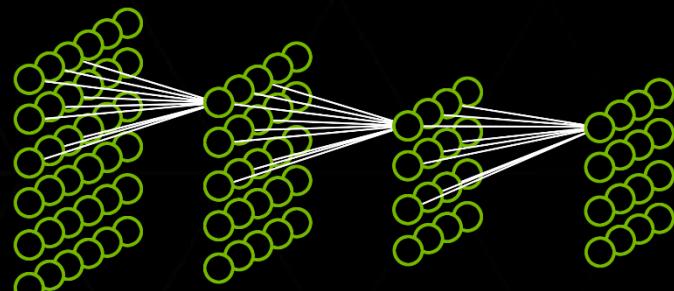
Синтез



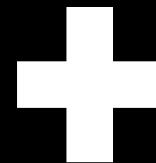
ТЕКСТУРНЫЕ ПРИЗНАКИ



Признаки



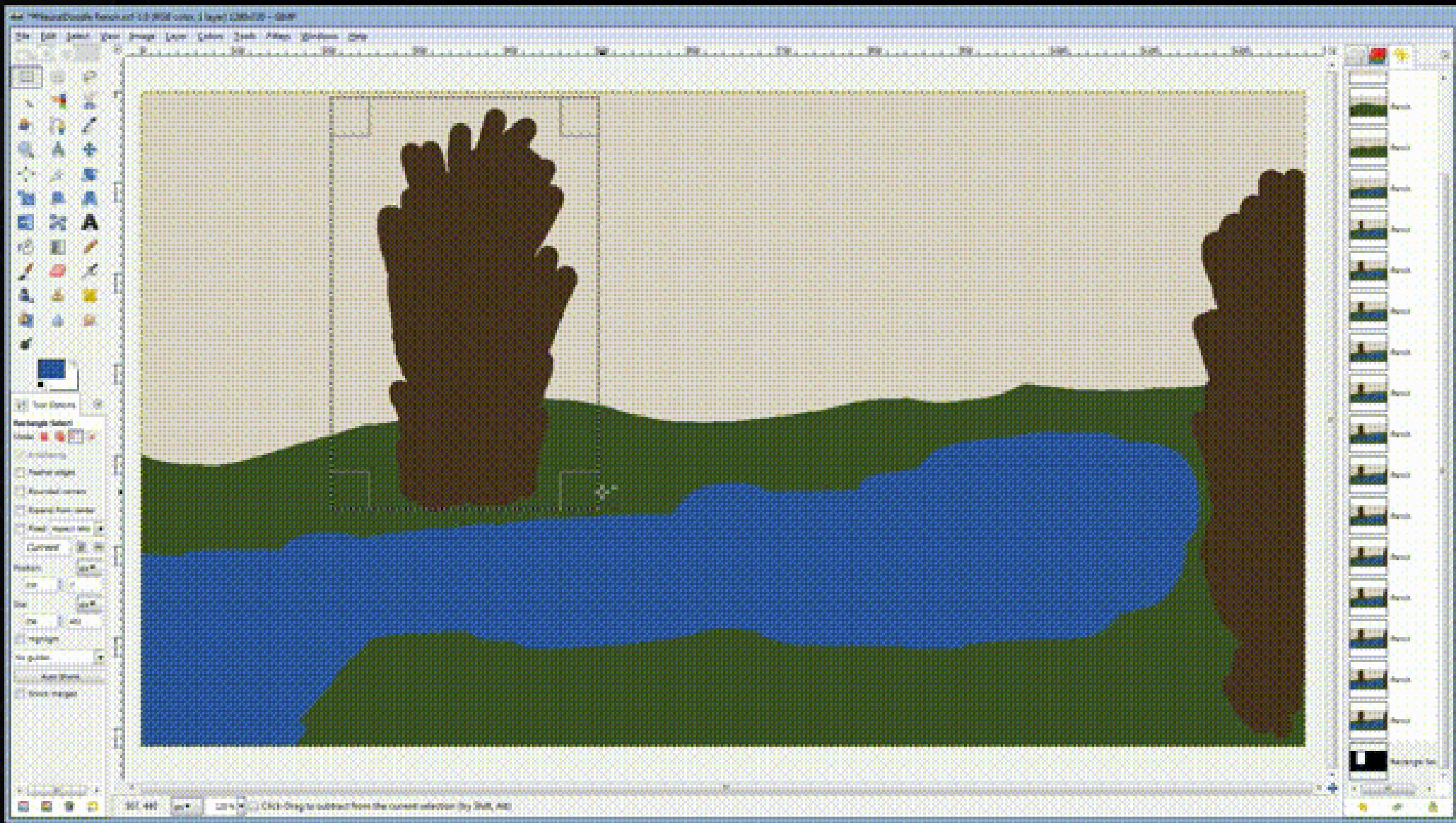
Текстурные
признаки



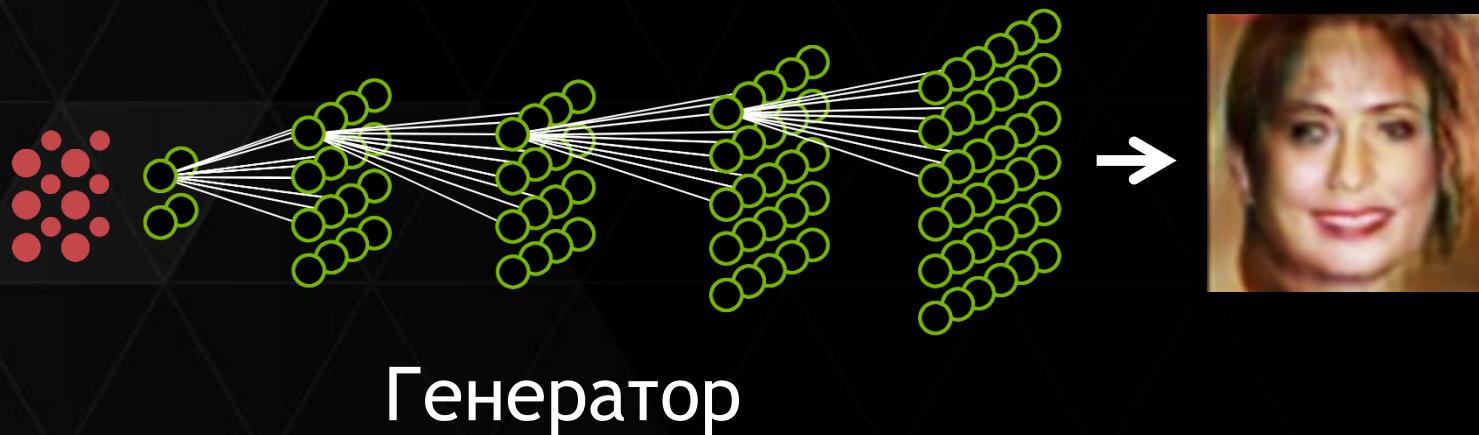
ПЕРЕНОС СТИЛЯ



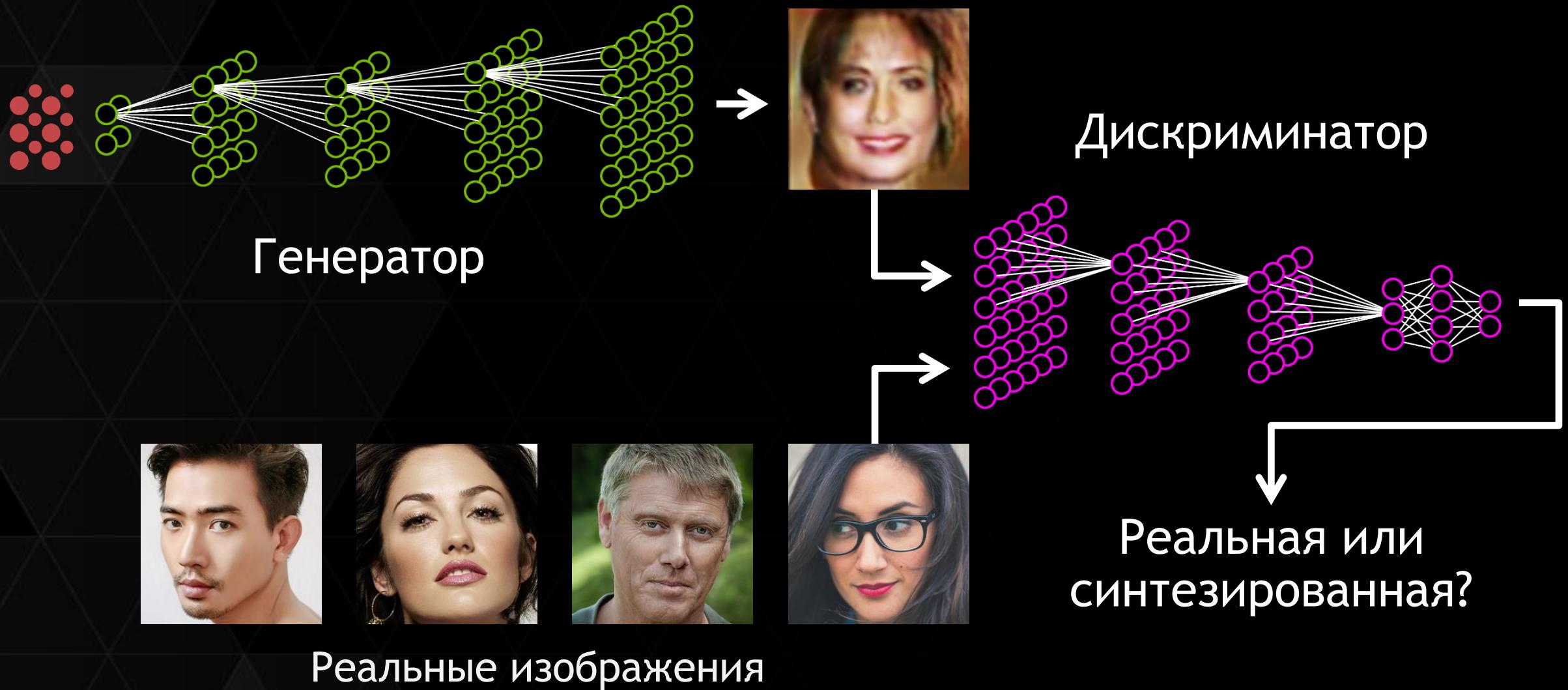
NEURAL DOODLE



СОСТАВАТЕЛЬНЫЕ СЕТИ (GAN)



СОСТАВАТЕЛЬНЫЕ СЕТИ (GAN)

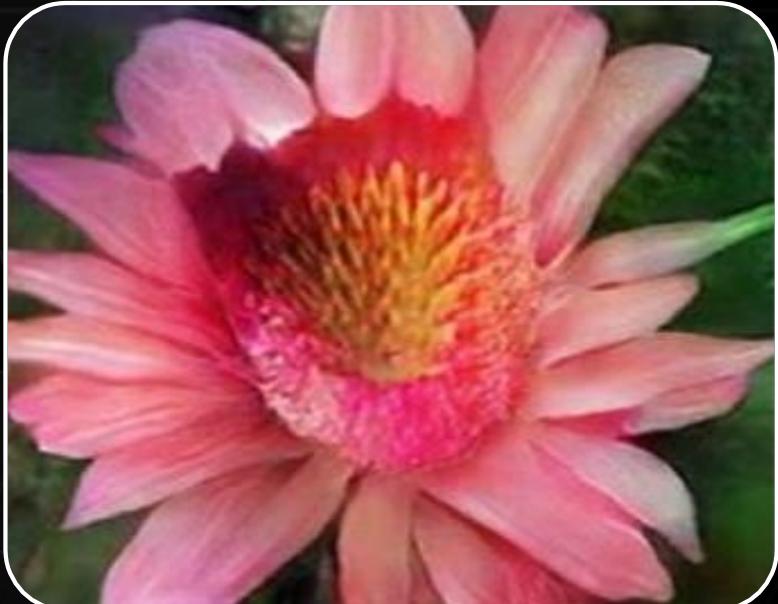


СИНТЕЗ ИЗОБРАЖЕНИЙ



ТЕКСТ → ИЗОБРАЖЕНИЕ

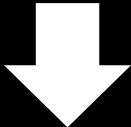
This flower has overlapping pink pointed petals surrounding a ring of short yellow filaments



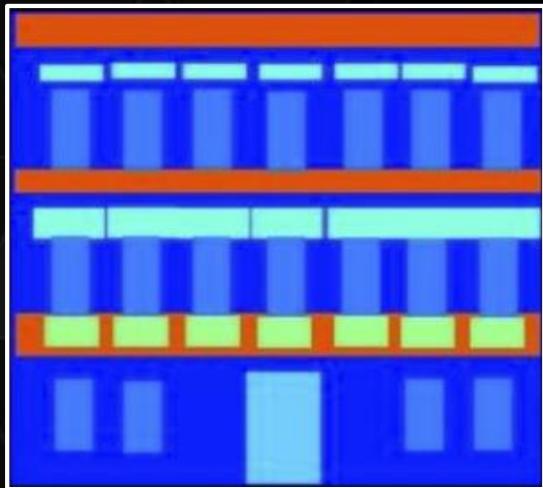
This bird is white with some black on its head and wings, and has a long orange beak



This bird has a yellow belly and tarsus, grey back, wings, and brown throat, nape with a black face

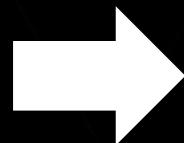
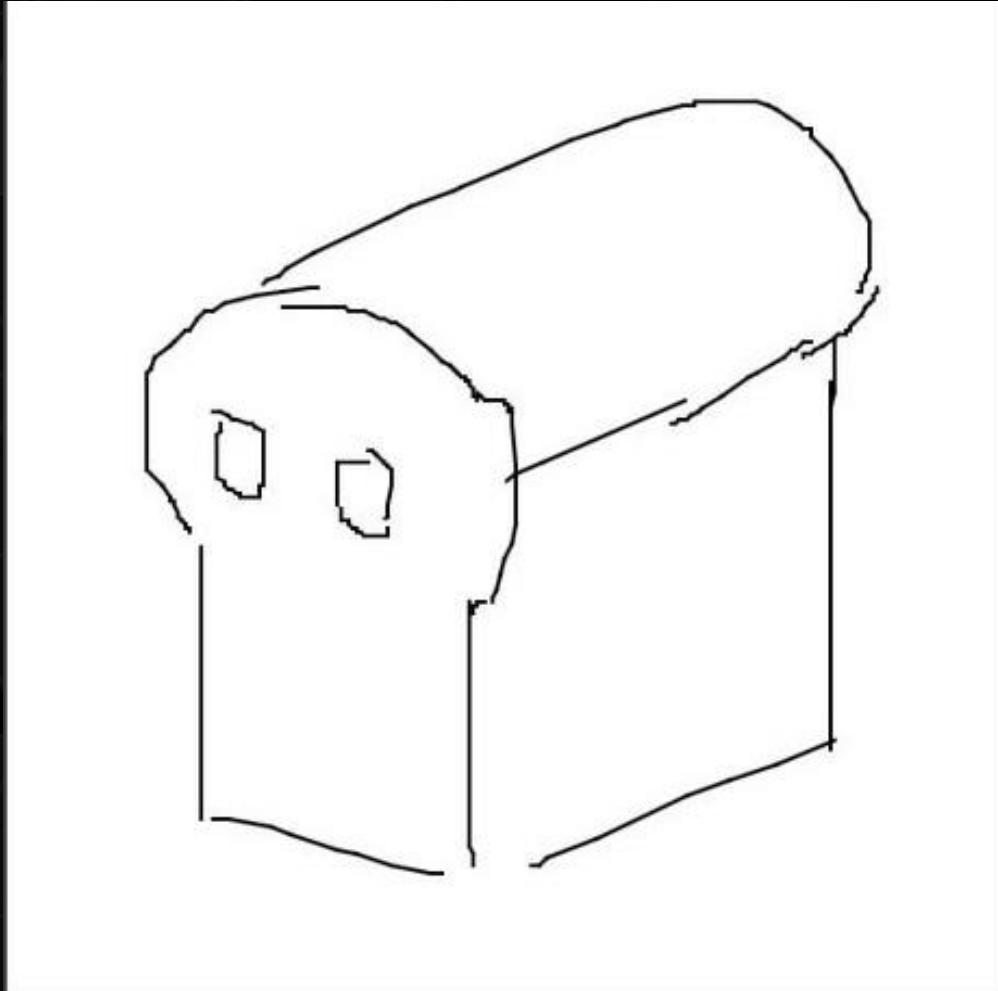


ИЗОБРАЖЕНИЕ → ИЗОБРАЖЕНИЕ

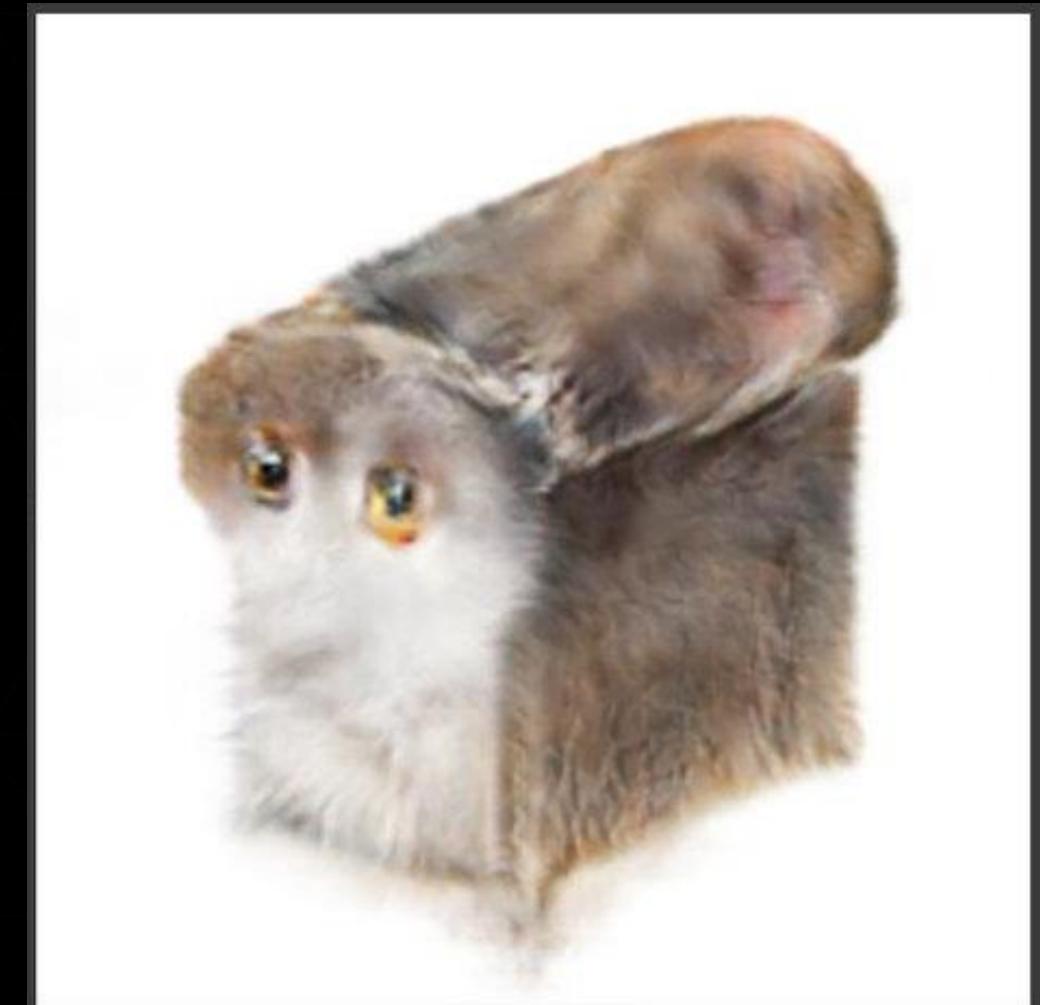


ИЗОБРАЖЕНИЕ → ИЗОБРАЖЕНИЕ

Вход

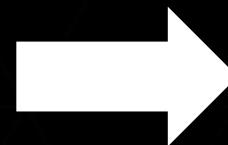


Выход



ВИДЕО → ВИДЕО

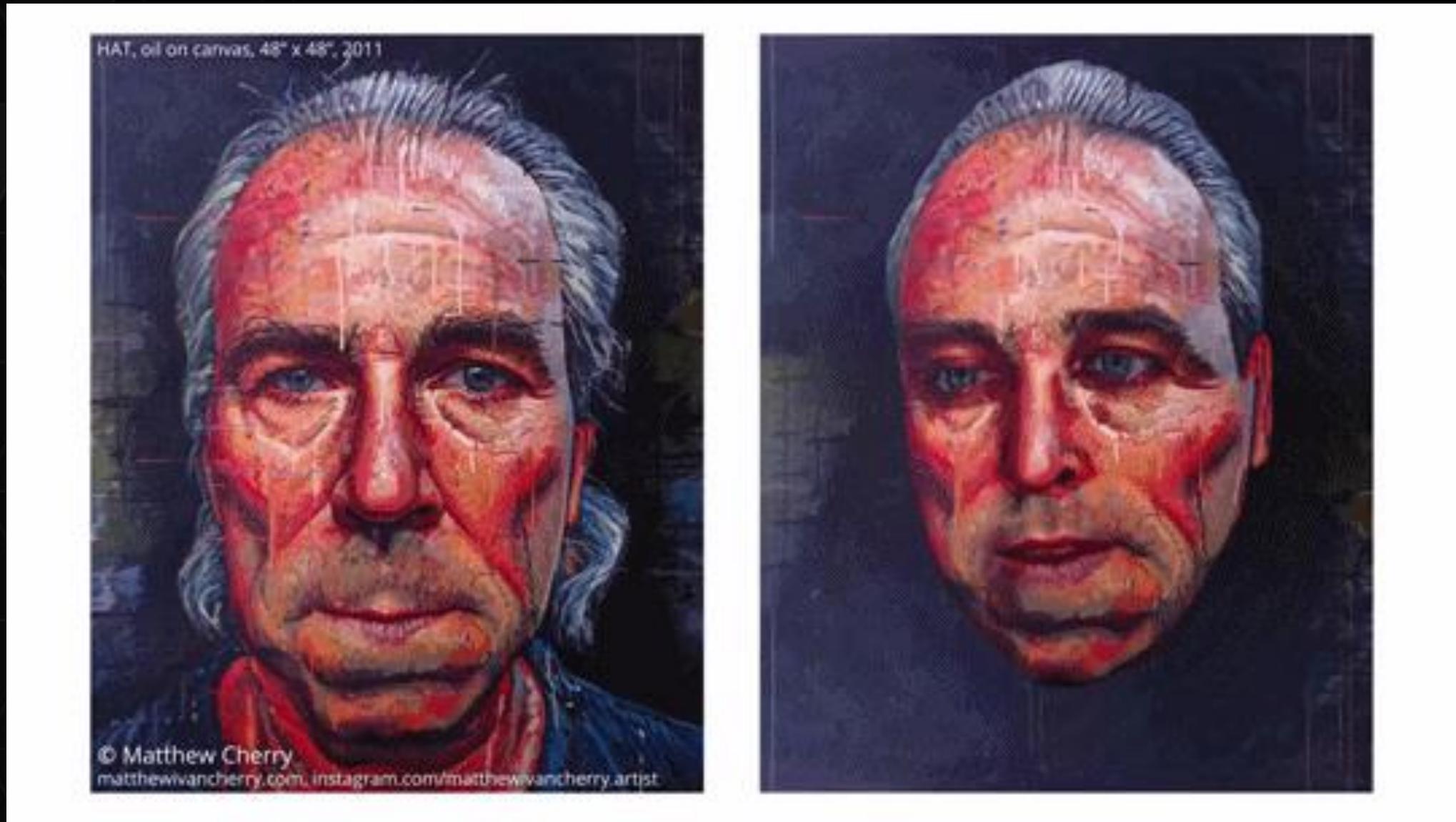
Вход



Выход



ВИДЕО → ВИДЕО



© Matthew Cherry
matthewivancherry.com, instagram.com/matthewivancherry.artist

ТЕКСТ → АУДИО



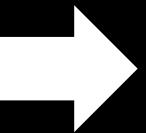
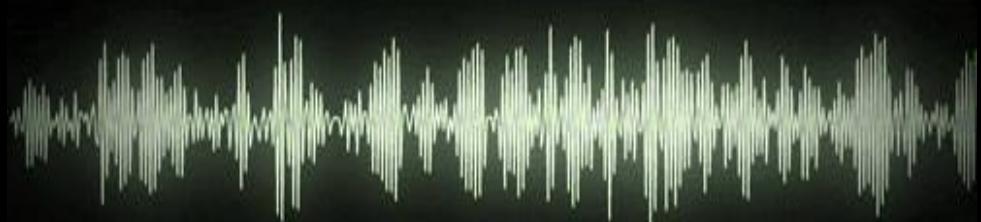
“Сегодня отличная погода, не правда ли?”



WaveNet



АУДИО → ВИДЕО

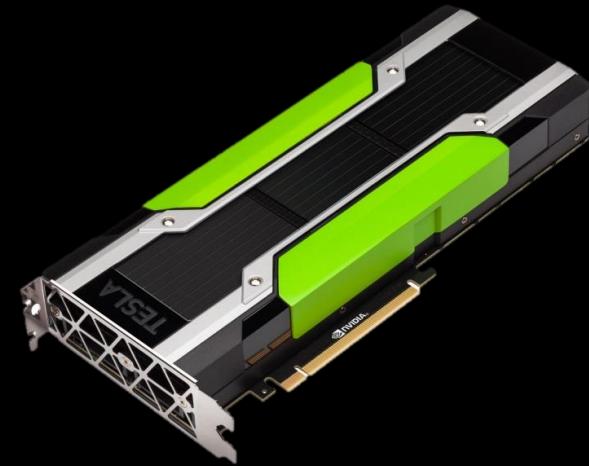
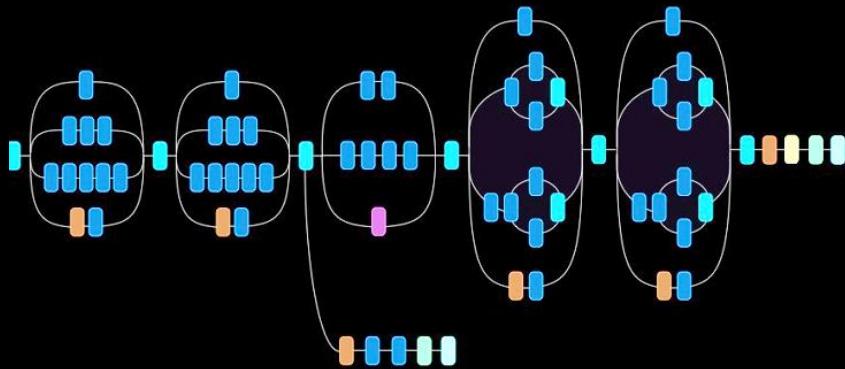


Запись голоса



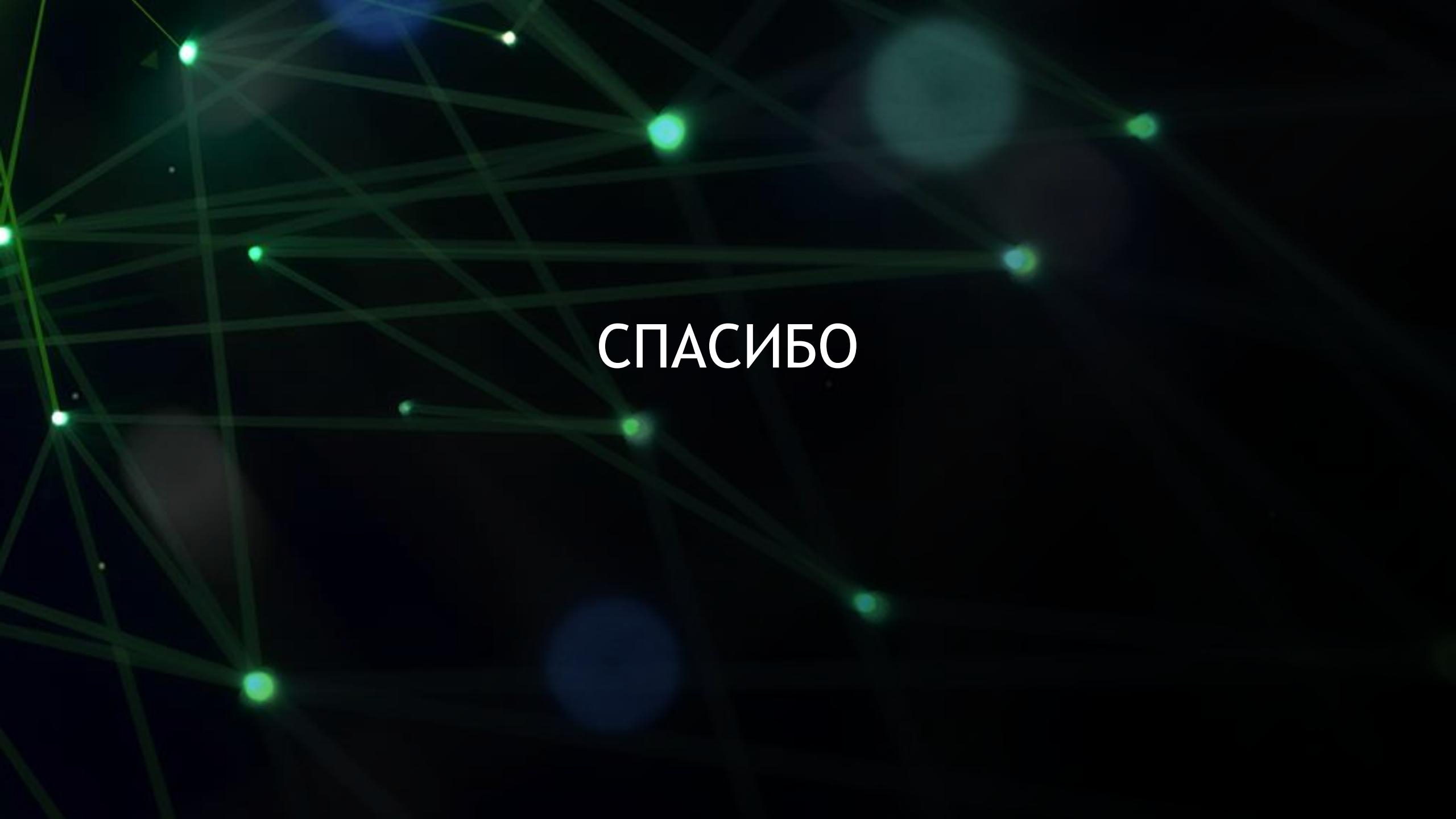
ПРИЧИНЫ УСПЕХА ГЛУБОКОГО ОБУЧЕНИЯ

- ✓ Совершенствующиеся алгоритмы и архитектуры
- ✓ Доступные объемы данных
- ✓ Ускорение обучения и вывода с помощью GPU



РЕЗЮМЕ

- Решаемые задачи: **отображение** X в Y:
 - Текст, изображения, видео, звук, низкоразмерные данные
- Типы архитектур:
 - Для распознавания и синтеза визуальных данных
 - **свёрточные** сети (**CNN**)
 - Для распознавания и синтеза последовательностей
 - **рекуррентные** сети (**RNN**)
 - Для улучшения качества синтеза
 - **состязательные** сети (**GAN**)
- Факторы успеха: топология, данные, железо (GPU)



СПАСИБО