

Глубинное обучение

Лекция 1

Введение в глубинное обучение

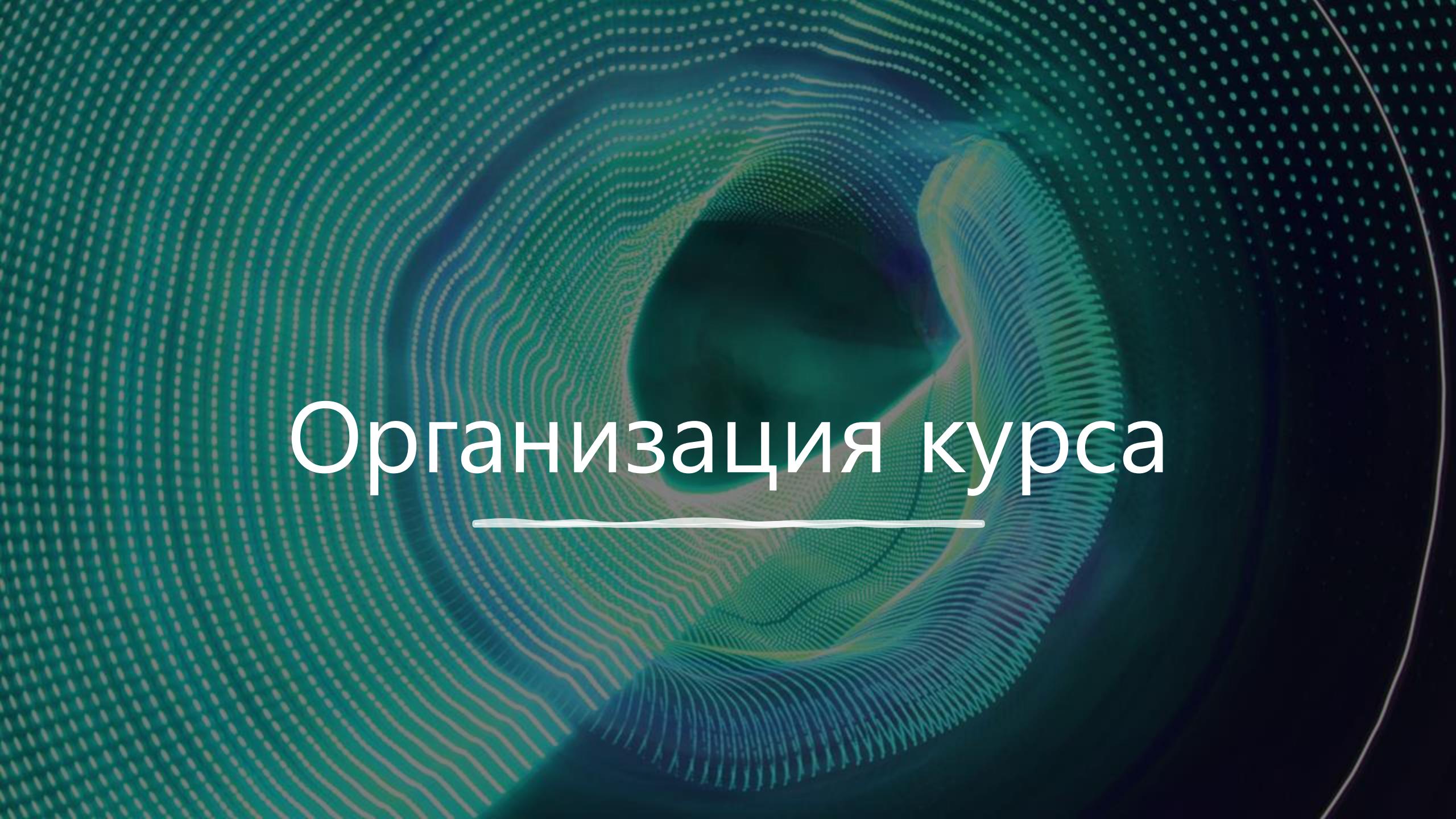
Михаил Гущин

mhushchyn@hse.ru

НИУ ВШЭ, 2025



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ



Организация курса

Наша команда

- ▶ Лектор:
 - Михаил Гущин
- ▶ Семинаристы:
 - Азиз Тёмирханов
- ▶ Ассистенты:
 - Анастасия Дорофеева
 - Артём Мантуров

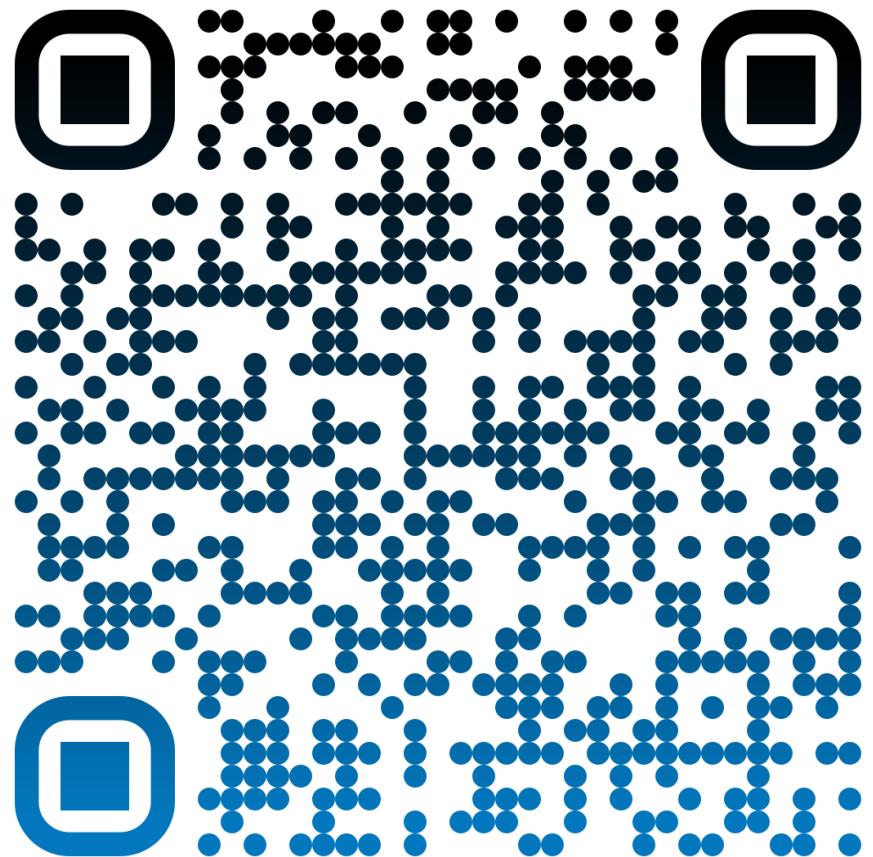
Чат курса в ТГ

- ▶ Важные объявления
- ▶ Можно задавать вопросы
- ▶ Дополнительные материалы



Вики курса

- ▶ Материалы лекций и семинаров
- ▶ Домашние задания
- ▶ Другая полезная информация по курсу
- ▶ Ссылки на курсы и книги



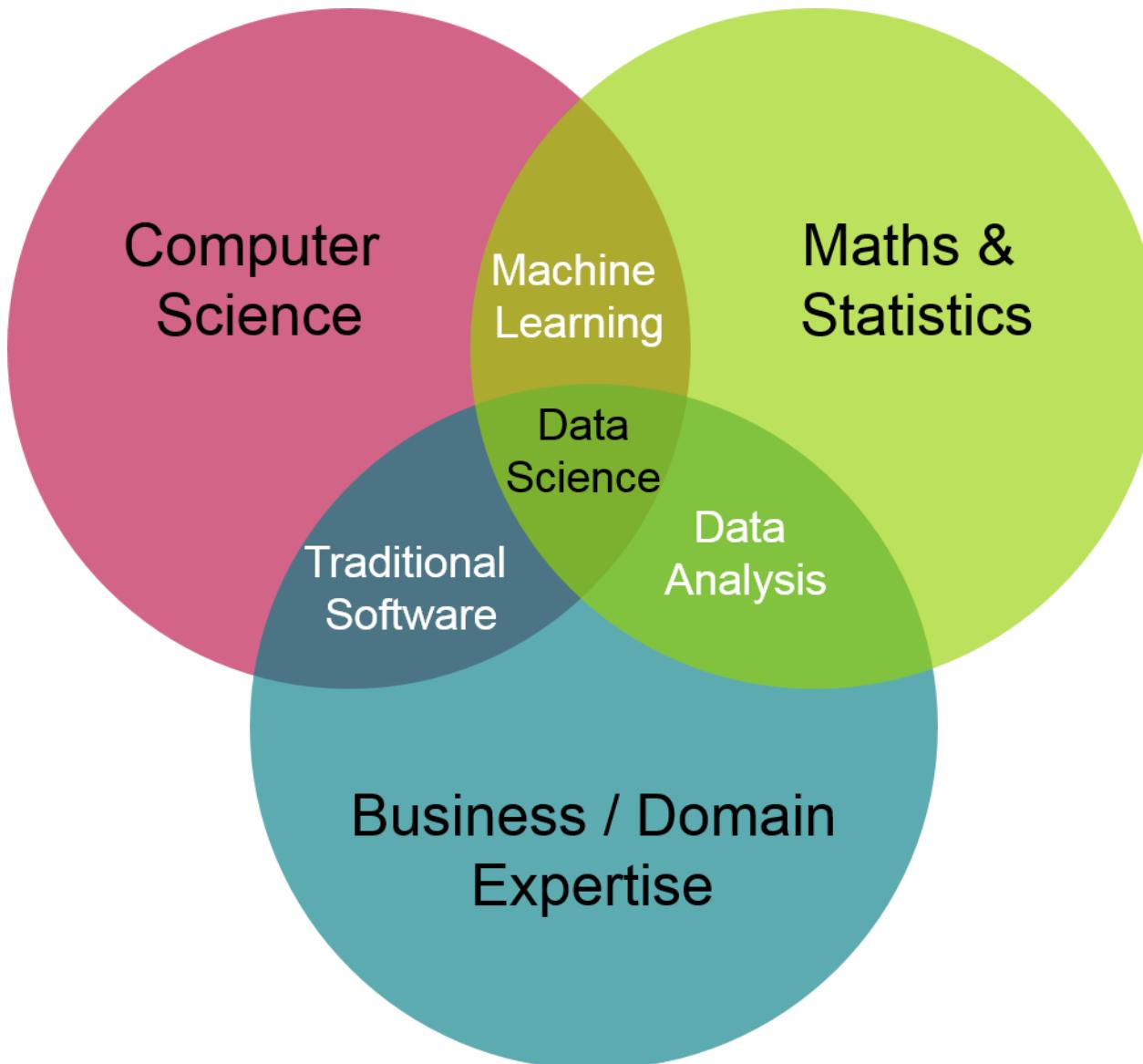
Занятия

- ▶ Лекции
 - по вторникам в 18:10 в G115
- ▶ Семинары
 - по вторникам в 19:40 в G115



Оценки

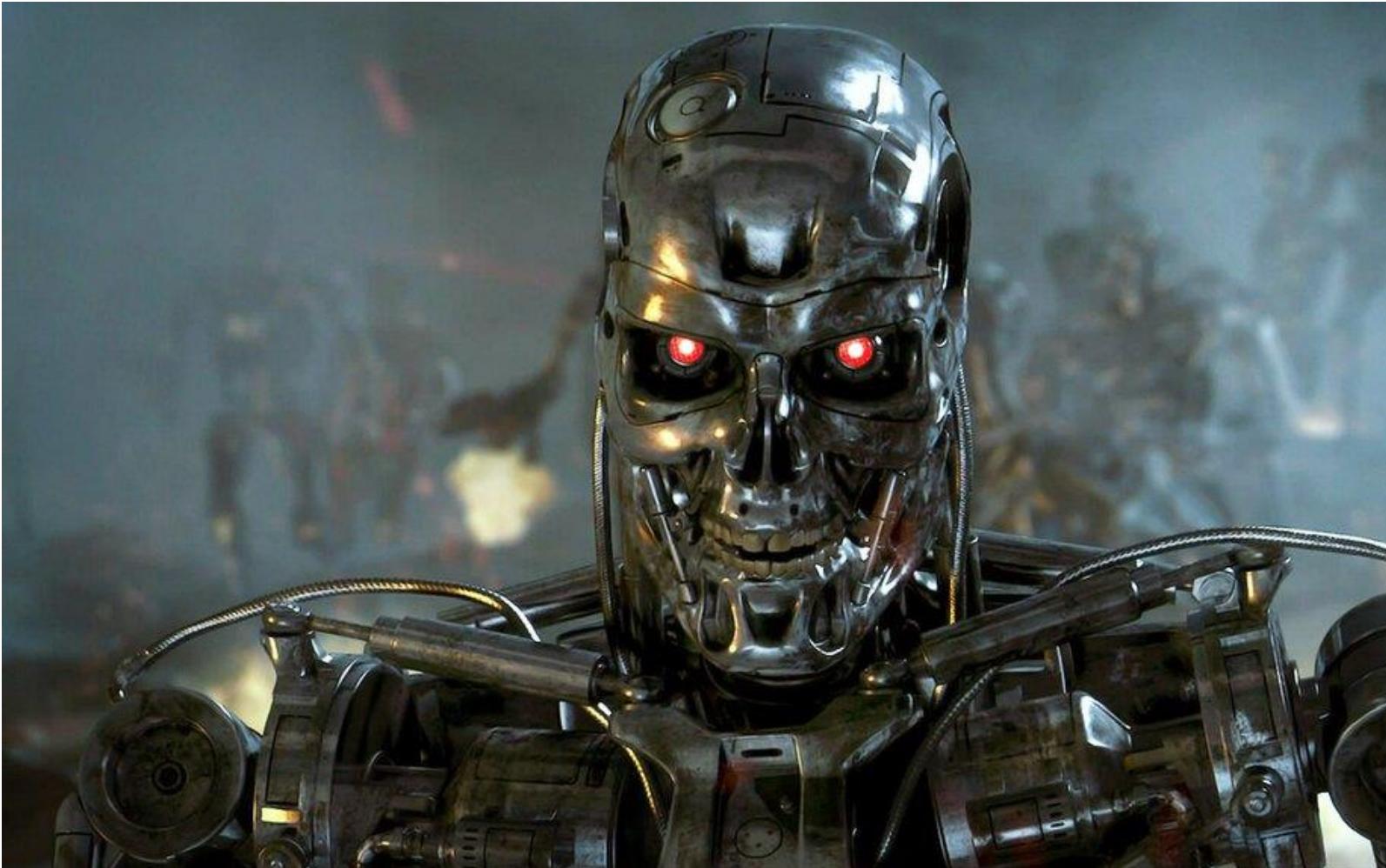
- ▶ Домашние задания
 - Практические задачи на программирование
 - Теоретические задачи
 - Мягкий и жесткий дедлайны
- ▶ Контрольная работа в ноябре
 - Теоретические вопросы + задачи
- ▶ Экзамен в конце курса
 - Теоретические вопросы + задачи
- ▶ Правила выставления оценок:
$$\text{Итоговая} = \text{Округление} (0.5 * \text{ДЗ} + 0.2 * \text{КР} + 0.3 * \text{Э})$$



ИИ в кино



Терминатор



**В КЕМБРИДЖЕ ИСКУССТВЕННЫЙ
ИНТЕЛЛЕКТ ВСТУПИЛ В ДЕБАТЫ С ЛЮДЬМИ
И УБЕДИЛ ИХ, ЧТО ОН НЕ ОПАСЕН.**



Ссылка на новость: <https://habr.com/ru/companies/pochtoy/articles/477540>

Я, робот



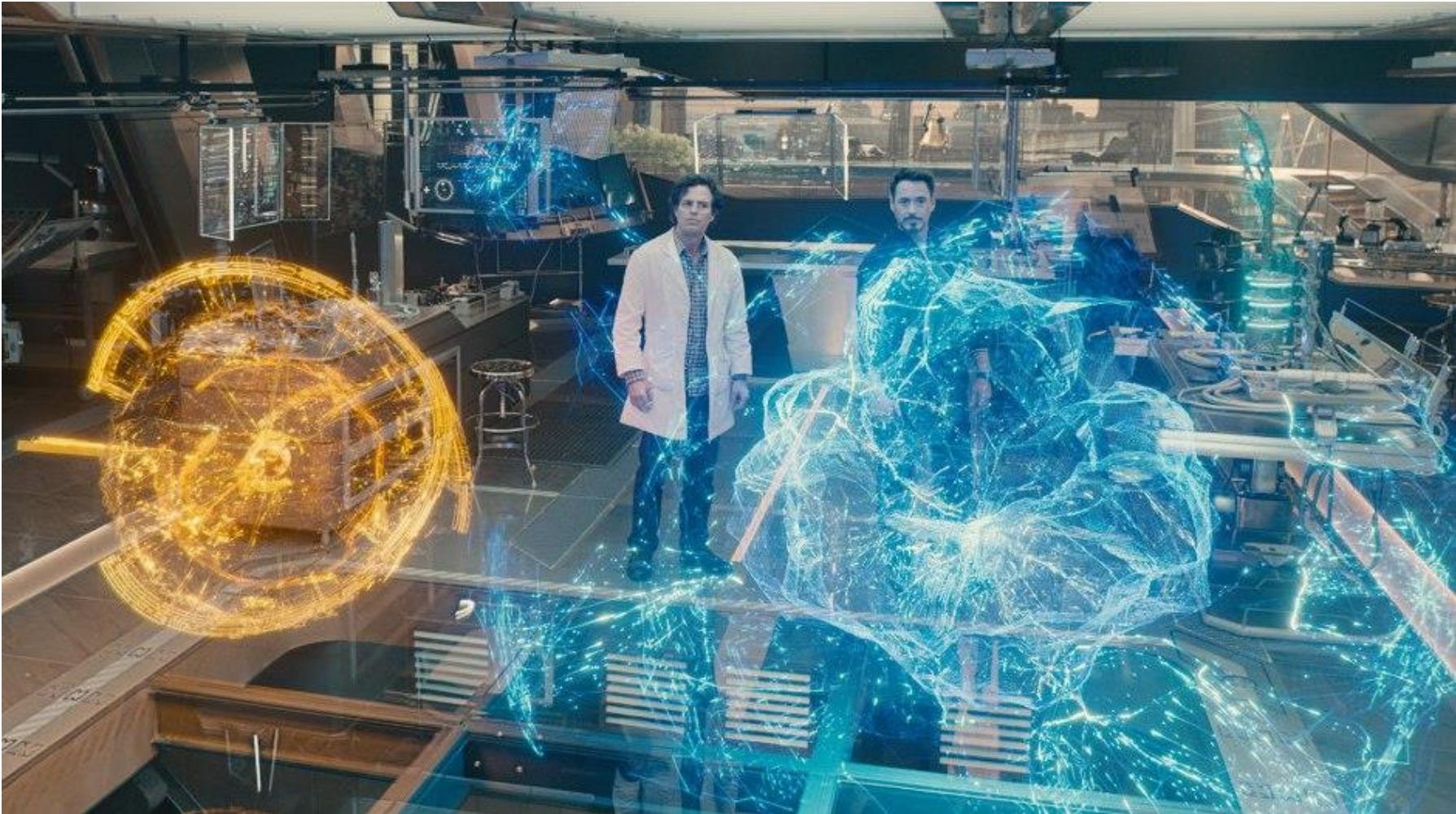
ВАЛЛ-И



Превосходство



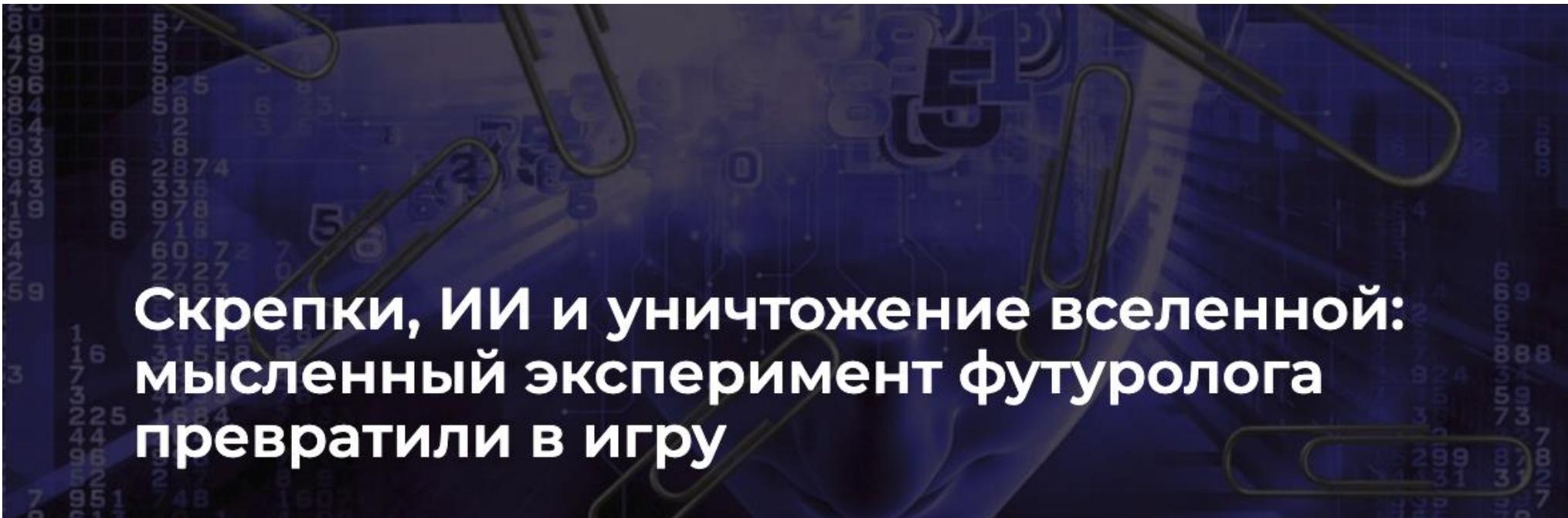
Железный человек / Мстители



Индекс страха



ИИ и скрепки

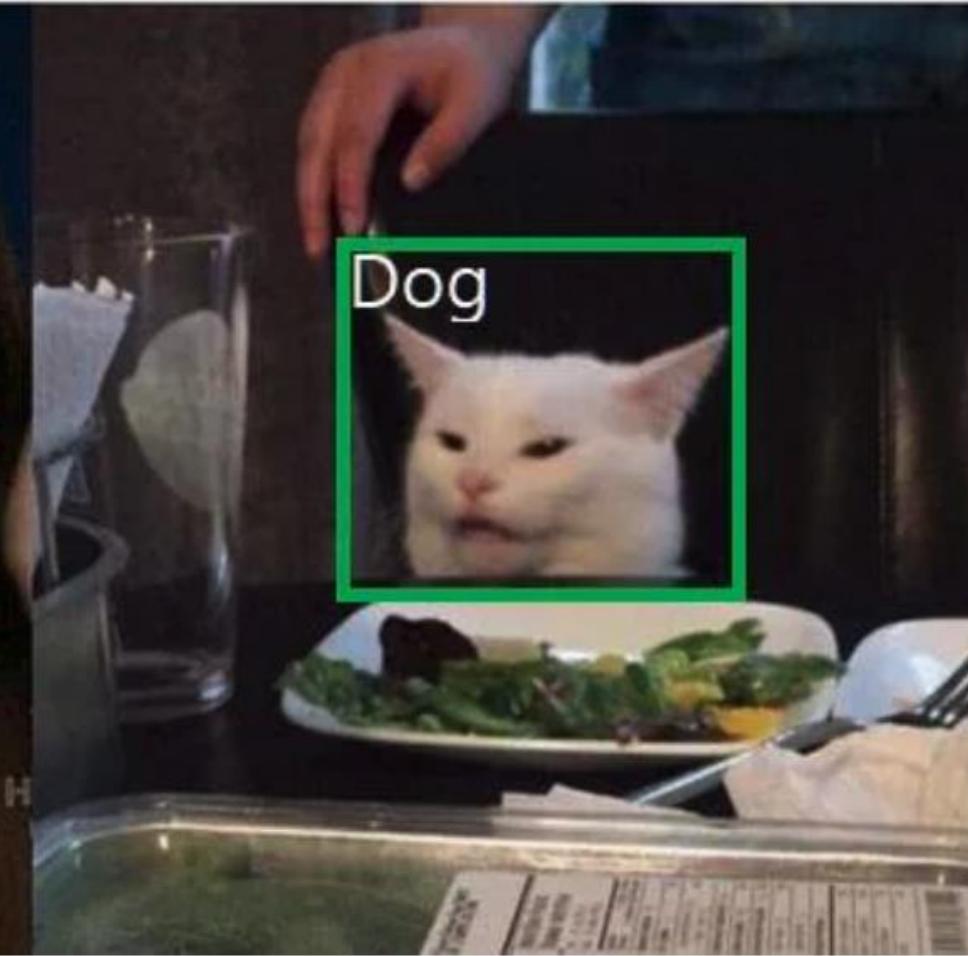


- ▶ Статья: <https://rb.ru/story/universal-paperclips/>
- ▶ Игра-кликер: <http://www.decisionproblem.com/paperclips/index2.html>

People with no idea
about AI, telling me my
AI will destroy the world



Me wondering why my
neural network is
classifying a cat as a dog..



Реальные приложения



Чатботы

- Example-1: Text → Text+Image+Audio

I've recently started practicing mindfulness and meditation as part of my daily routine, but sometimes it's challenging. Can you help me with some techniques?

Of course! Mindfulness and meditation can be a wonderful way to cultivate a sense of calm and focus. One technique you might find helpful is called "mindful breathing." It involves taking slow, deep breaths and focusing your attention on the sensation of the breath entering and leaving your body. Another technique is "body scan meditation," where you consciously scan your entire body from head to toe, paying attention to any sensations or discomfort. You can also try guided meditations that provide a structured practice. Would you like me to find a video that demonstrates mindful breathing or body scan meditation?

I wish to listen to a music for the meditation first.

I guess you may want some musics that sound soft.

▶ 0:00 / 0:09

Base on the audio content, visualize it in an image.



Can you infer what time is in the picture?

Based on the image, it is likely that the time is either early morning or late afternoon. The sunlight is either low or setting, creating a serene atmosphere.

Чатботы

How many cars are in this picture?



There are two cars shown in this picture.

Tell me what kind of car are they?

The two cars in the picture are SUVs.

Can you show me the sound of starting the engines of them.

Sure, here's the sound of the engine of the SUV:

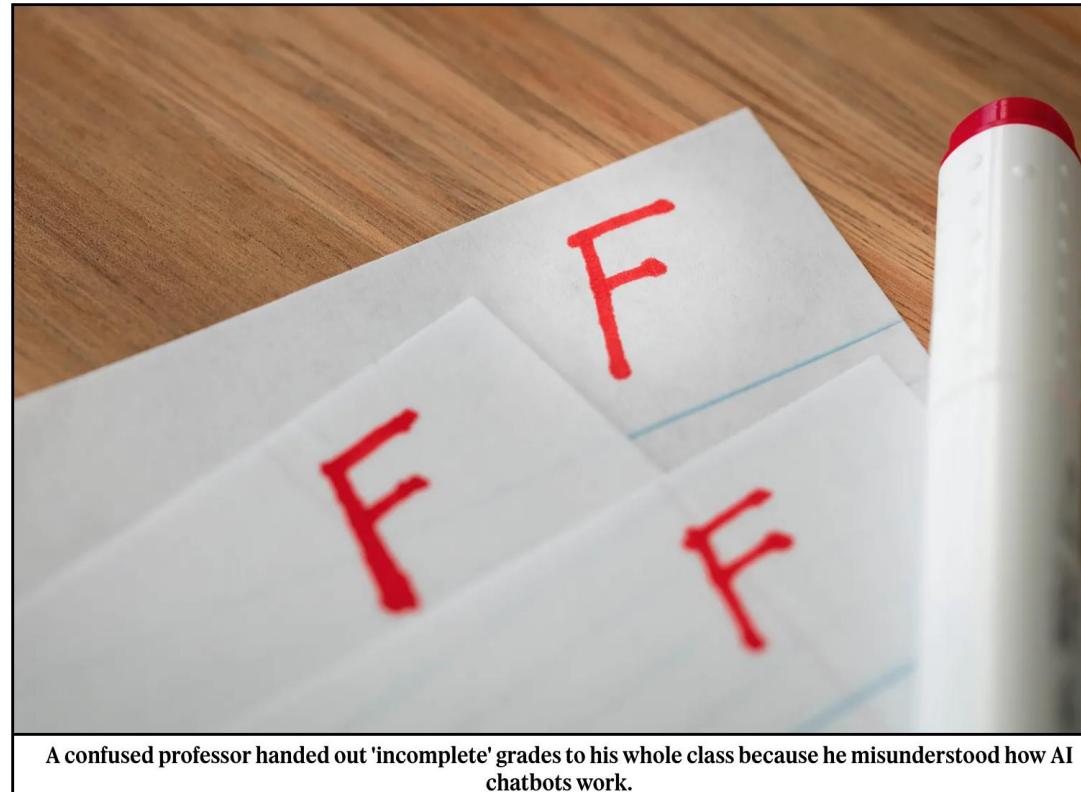
▶ 0:00 / 0:09

PROFESSOR FLUNKS ALL HIS STUDENTS AFTER CHATGPT FALSELY CLAIMS IT WROTE THEIR PAPERS

Texas A&M University–Commerce seniors who have already graduated were denied their diplomas because of an instructor who incorrectly used AI software to detect cheating

By MILES KLEE

MAY 17, 2023

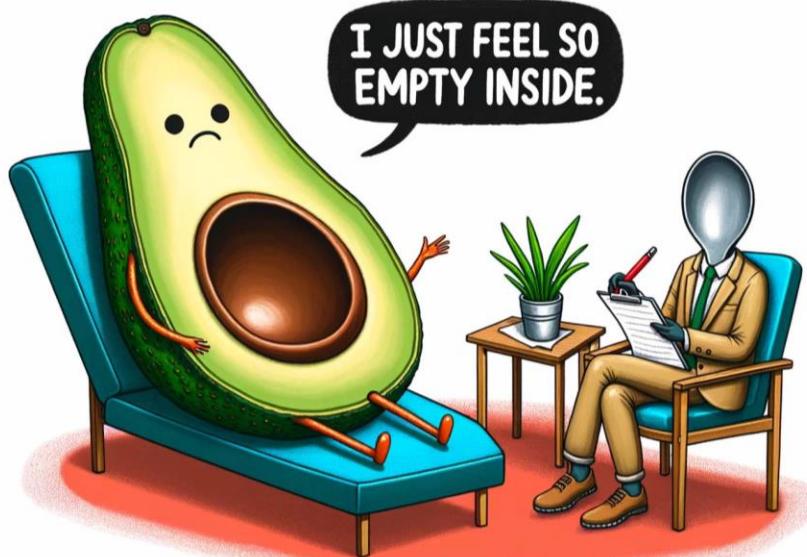


Генерация фотографий



Source: Goodfellow et al., 2014; Radford et al., 2016; Liu & Tuzel, 2016;
Karras et al., 2018; Karras et al., 2019; Goodfellow, 2019; Karras et al.,
2020; AI Index, 2021; Vahdat et al., 2021

Генерация изображений (Dall-E 3)

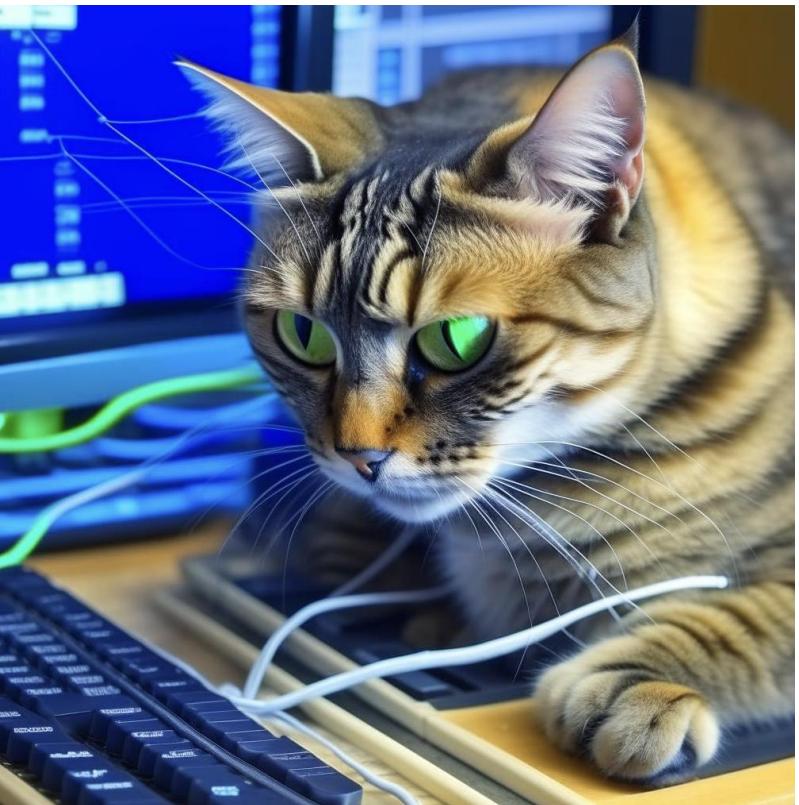


An illustration of an avocado sitting in a therapist's chair, saying 'I just feel so empty inside' with a pit-sized hole in its center. The therapist, a spoon, scribbles notes



Tiny potato kings wearing majestic crowns, sitting on thrones, overseeing their vast potato kingdom filled with potato subjects and potato castles.

Генерация изображений



Кот фиксит баг в обучении
нейронной сети



Розовый фламинго стоит на
одной ноге в воде

Ссылка: <https://www.sberbank.com/promo/kandinsky>

Генерация и видео

ИИ Редактор для Создания Видео



Сцены с несколькими аватарами

Имитирование взаимодействия между аватарами - идеальное решение для обучающих видео и образовательного контента.



Примечания и Эффекты

Добавляйте заметки в сценарий, чтобы запустить анимацию и привлечь внимание аудитории к ключевой информации.



Управление Движениями Аватарами

Задавайте движения и жесты аватара, чтобы привлечь внимание к основным моментам в вашем контенте

Перевод Видео на 80 Языков



Перевод Текста в Речь на Несколько Языков

Создавайте закадровые голоса с искусственным интеллектом на более чем 80 языках, используя наши голоса искусственного интеллекта или свой собственный голосовой клон.



Видеодубляж

Загрузите видео и получите его перевод с идеальной синхронизацией губ.



Автоматизированный Перевод Видео

По мере создания автоматически переводите видеосценарий и текст на экране.

www.aistudios.com

Как отличить картинки



Erkhyan
@erkhyan@yiff.life

◀ prg_memes

Что забавно: если хочешь понять,
сгенерирован ли рисунок нейросеткой,
следуй правилам, по которым в старых
сказках распознавали нечистую силу.

«Посчитай пальцы, посчитай костяшки,
посчитай зубы, проверь тень...»
...и НИ ПРИ КАКИХ обстоятельствах не
соглашайся на сделку с этим отродьем.

Deepfake



John Wick 4: Dangerous Cleanup
7,4 млн просмотров



This is you right now 😂
1,2 млн просмотров



Katana wins
18 млн просмотров



Thank you, my followers ❤️
1,9 млн просмотров



10 minutes of eternity...
1,9 млн просмотров



Keanu Reeves lives with his girlfriend
19 млн просмотров



Did you recognize everyone?
2,1 млн просмотров



When I came back from filming
1,2 млн просмотров



You spin USB right 'round, baby, right 'round... #keanu...
1,5 млн просмотров



Who is the best dancer? 😎
#keanu #reeves #dance #fyp
7,9 млн просмотров



Mirror trick 😜 #keanu #reeves #mirror #tenet
1,1 млн просмотров



Dressing up like a cool guy.
#reeves #keanu #dressup
1,1 млн просмотров

Deepfake



Unreal Keanu Reeves

@unreal_keanu

617 тыс. подписчиков

ГЛАВНАЯ

SHORTS

ПЛЕЙЛИСТЫ

СООБЩЕСТВО

КАНАЛЫ

О КАНАЛЕ

Описание

This is the official deepfake channel of actor Keanu Reeves.
Like, share, subscribe!

Дополнительно

Для коммерческих запросов:

[Показать адрес электронной почты](#)

Страна:

Грузия

Source: https://www.youtube.com/@unreal_keanu/about

Deepfake



«Женщина развелась с мужем-миллионером ради «Брэда Питта», которому отдала 850 тысяч долларов.»



r/nottheonion • 21 hr. ago

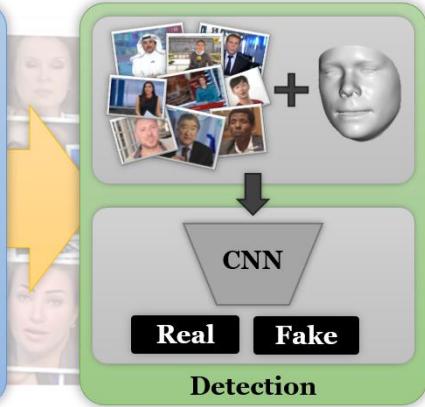
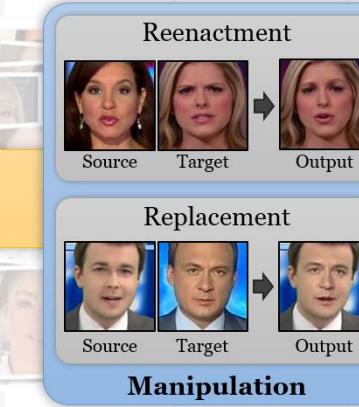
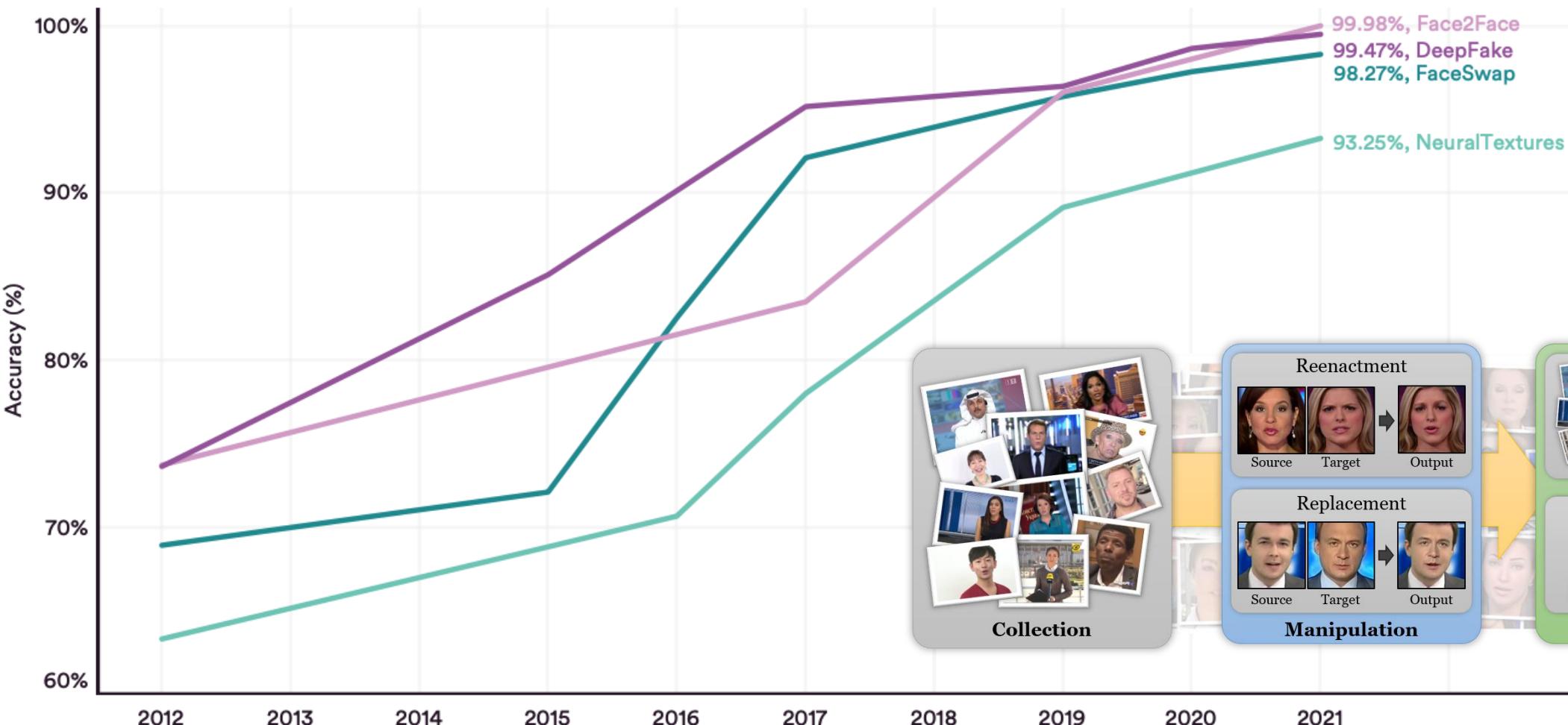
Woman who warned social media not to fall for Keanu Reeves romance scam becomes homeless after falling for Keanu Reeves romance scam



Deepfake detection

FACEFORENSICS++: ACCURACY

Source: arXiv, 2021 | Chart: 2022 AI Index Report

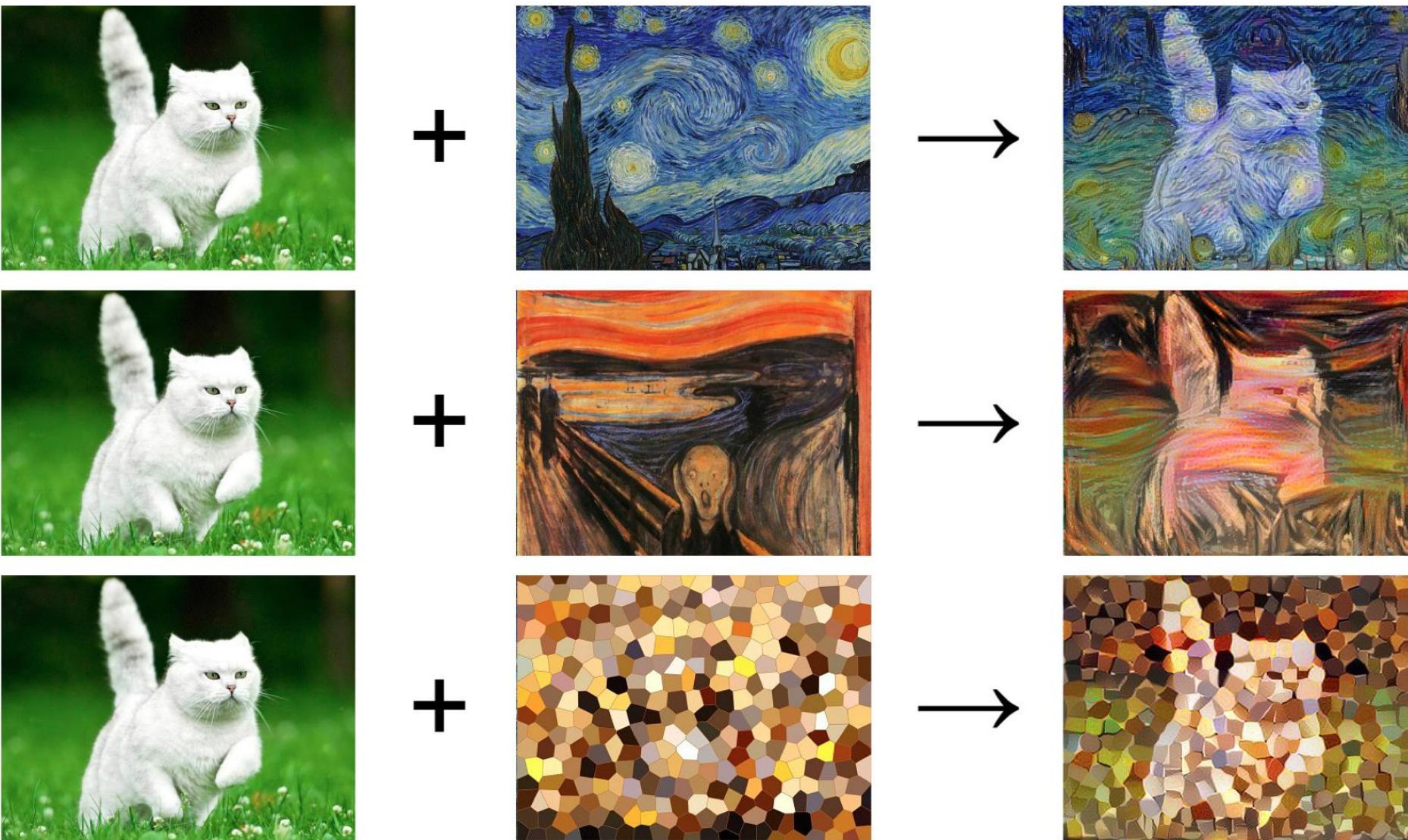




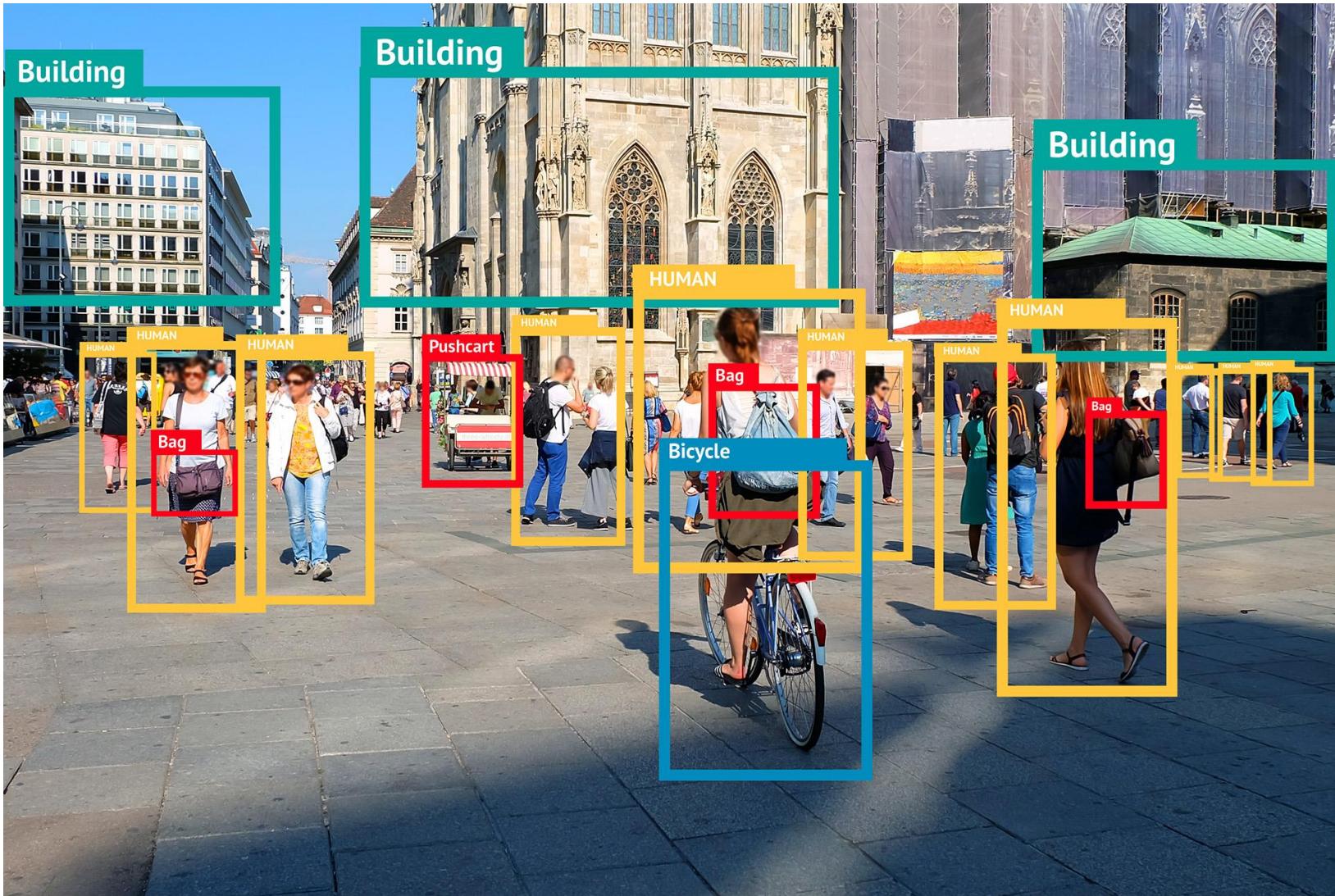
**Замечаешь всё
меньше картинок,
сгенерированных ИИ**

**Замечаешь всё
меньше картинок,
сгенерированных ИИ**

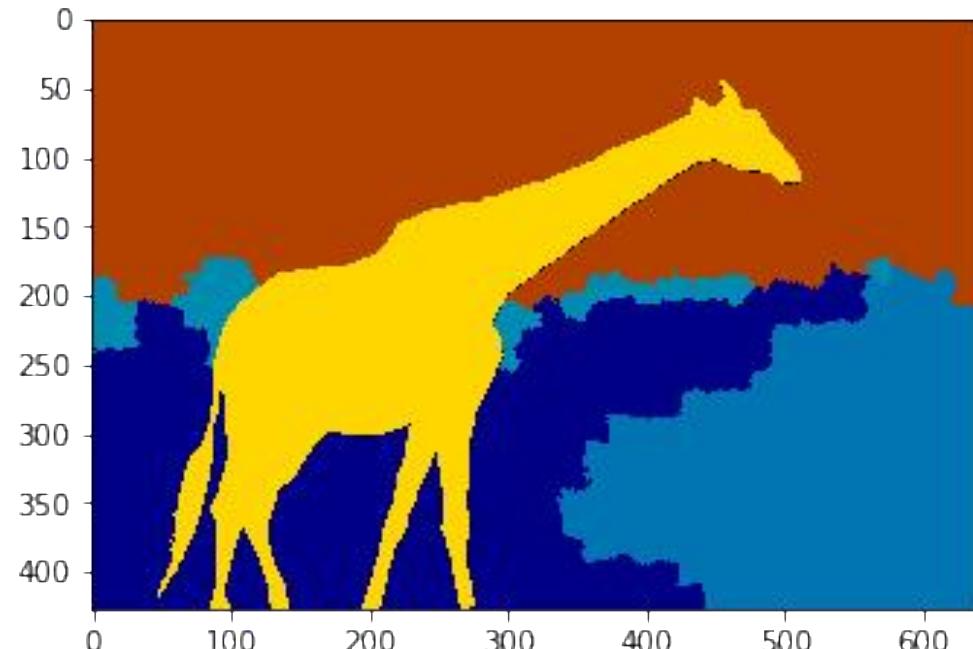
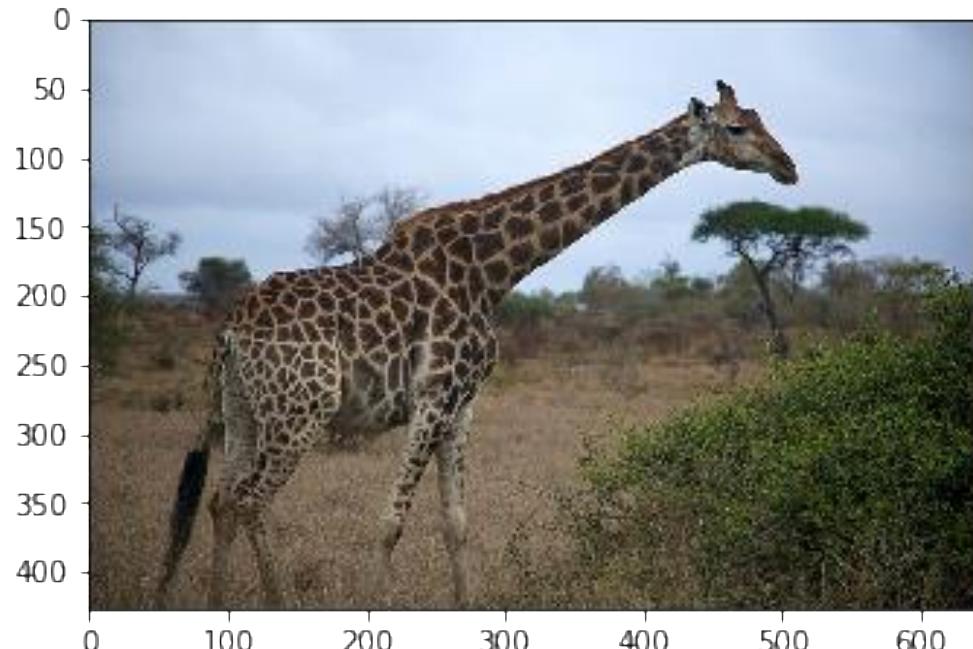
Перенос стиля изображения



Детектирование объектов



Сегментация изображений



Поиск информации

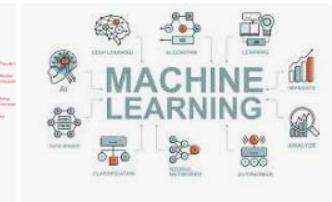
Google machine learning X | ⚡ 🔎

All Images Videos News Books More Tools

About 2,080,000,000 results (0.54 seconds)

Machine learning (ML) is a type of artificial intelligence (AI) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so. Machine learning algorithms use historical data as input to predict new output values.

<https://www.techtarget.com/searchenterpriseai/definition> :: [What Is Machine Learning and Why Is It Important? - TechTarget](#)



About featured snippets · Feedback

https://en.wikipedia.org/wiki/Machine_learning :: [Machine learning - Wikipedia](#)

Machine learning (ML) is a field of inquiry devoted to understanding and building methods that 'learn', that is, methods that leverage data to improve ...

[Machine Learning \(journal\)](#) · [Machine learning control](#) · [Active learning \(machine...](#)

See results about

machine learning Dictionary definition >

People also ask :

What is machine learning with example? ▾

What are the 3 types of learning in machine learning? ▾

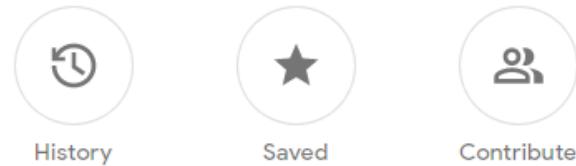
What is AI vs machine learning? ▾

Why is machine learning used? ▾

Feedback

Машинный перевод

The screenshot shows the Google Translate interface. At the top, there are three tabs: 'Text' (selected), 'Documents', and 'Websites'. Below this is a language detection bar with 'DETECT LANGUAGE' followed by 'ENGLISH', 'GERMAN', 'GEORGIAN', and a dropdown arrow. To the right is another language bar with 'GERMAN' (selected), 'ENGLISH', 'UKRAINIAN', and a dropdown arrow. The main content area displays a comparison between English and German translations of a text about machine translation tools. The English input is 'Machine Translation Tools: Translate Content with AI Software in 2022' and the German output is 'Tools für maschinelle Übersetzung: Übersetzen Sie Inhalte mit KI-Software im Jahr 2022'. Both versions have a green circular icon with a 'G'. Below the text are audio icons and a progress bar showing '69 / 5,000'. On the right side of the German translation are icons for a reply, a thumbs up, and a share. At the bottom right of the main area is a link 'Send feedback'. A small 'G' logo is also visible near the German translation.



Такси (self driving car)



waymo.com

Машинный перевод

Яндекс Браузер

Яндекс.Браузер обновился. Версия 21.8.2

1

Hi! I'm David and I lead the NLP team at Yandex

Закадровый перевод видео с английского

Нейросети Яндекса научились сами переводить и озвучивать видео на английском языке. Пока — не везде, но уже скоро любой ролик на английском можно будет смотреть на русском.

Сразу попробовать новую функцию можно [по ссылке](#).

Как включить перевод видео?

Подписывайтесь на новости Яндекс.Браузера:

Дзен ВКонтакте Твиттер Телеграм

Перевод не доступен для видео, у которых есть технические средства защиты авторских прав (DRM).

Голосовые помощники

The screenshot shows the Yandex.Alice mobile application interface. At the top center is the Alice logo, a stylized white egg inside a purple circle. Below it, a large white speech bubble contains the text "Привет, я Алиса!". The background is a solid purple color. In the center, the text "Я готова помочь" is displayed. Below this, there are six horizontal cards, each with an icon and text: 1. "Определить песню" (Identify song) with a musical note icon. 2. "Узнать, что на фото" (Know what's in the photo) with a camera icon. 3. "Включить сказку" (Turn on story) with a book icon. 4. "Одеться по погоде" (Dress according to the weather) with a coat icon. 5. "Поиграть" (Play) with a game controller icon. 6. "Построить маршрут" (Build route) with a map pin icon. 7. "Вызвать такси" (Call a taxi) with a car icon. 8. "Найти нужное место" (Find the right place) with a location pin icon. 9. "Купить на Беру" (Buy on Beru) with a shopping box icon.

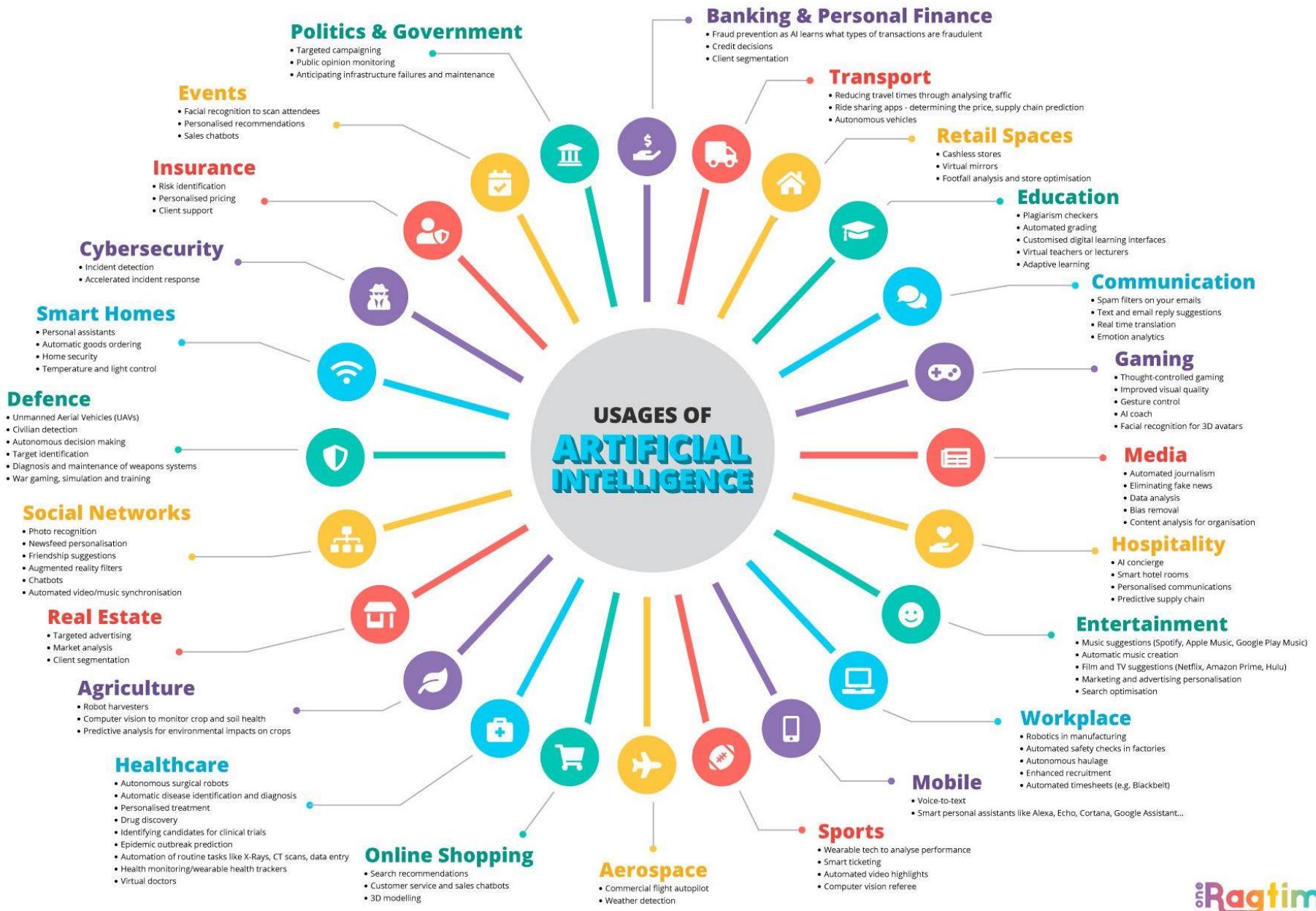
Привет, я Алиса!

Я готова помочь

Определить песню > Узнать, что на фото > Включить сказку >

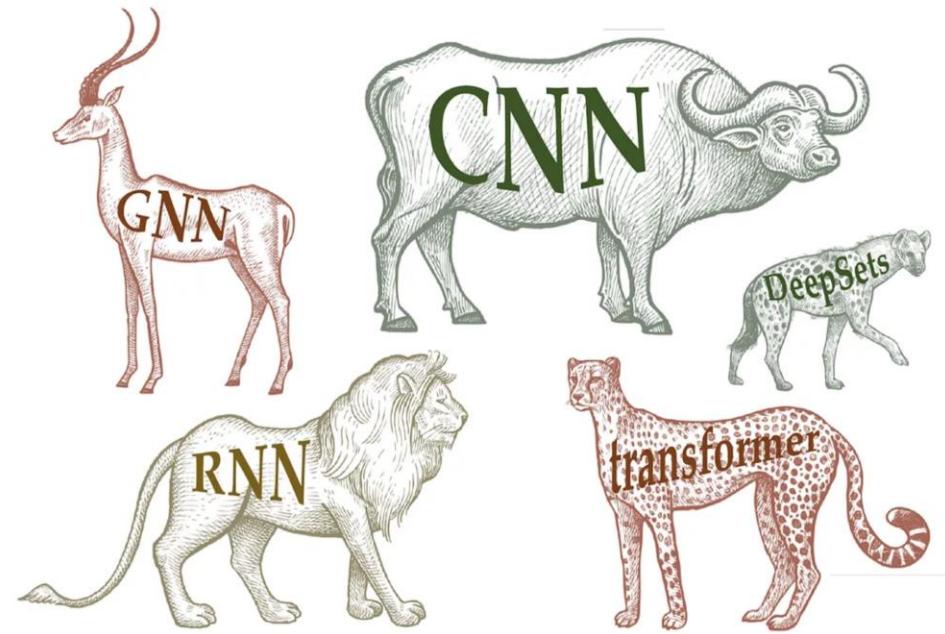
Одеться по погоде > Поиграть > Построить маршрут >

Вызвать такси > Найти нужное место > Купить на Беру >



В нашем курсе

- ▶ Полносвязные нейронные сети
- ▶ Сврточные нейронные сети
 - Классификация изображений
 - Детектирование объектов
 - Семантическая сегментация
- ▶ Рекуррентные нейронные сети
 - Анализ текстов
 - Генерация текста
- ▶ Трансформеры
 - Механизм внимания
 - Машинный перевод



Source: <https://geometricdeeplearning.com/lectures/>

Линейная регрессия. Аналитическое решение.

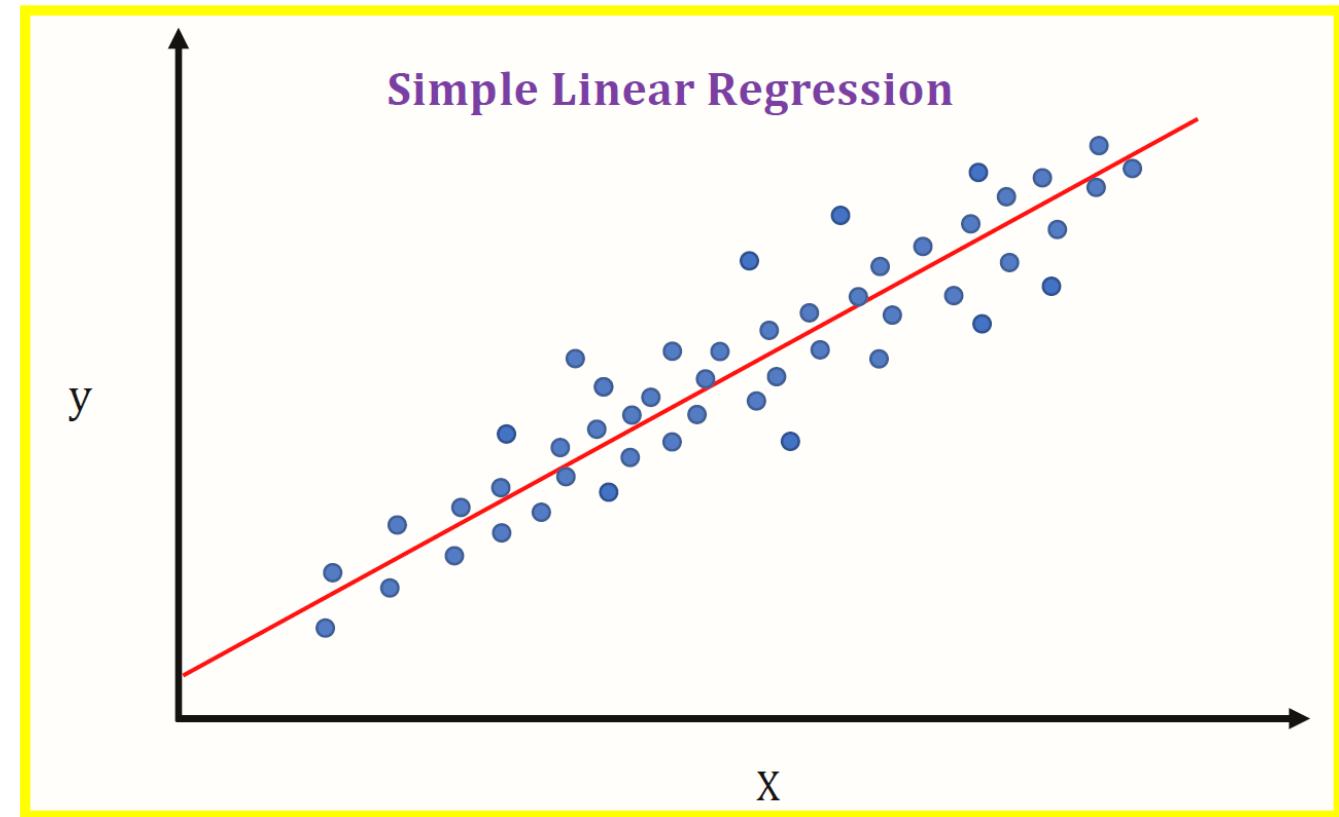
Примеры

- ▶ Прогноз продолжительности жизни
- ▶ Оценка рисков в банках
- ▶ Прогнозирование цены товара
- ▶ Прогнозирование объема продаж

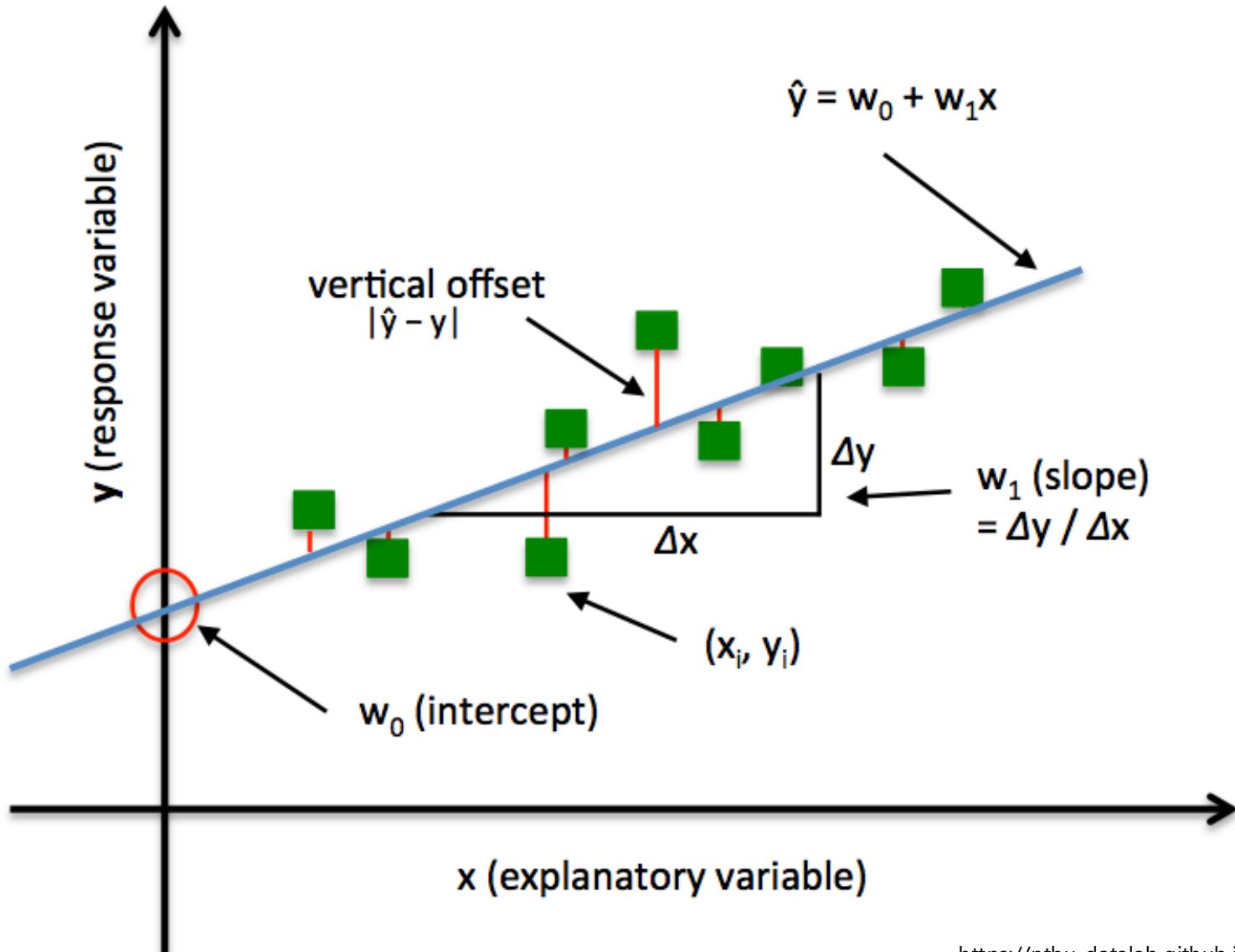


Задача регрессии

- ▶ Есть объекты (X)
- ▶ Нужно предсказать некоторую величину (y)
- ▶ Функция, которая описывает зависимость y от X - **модель регрессии**



Линейная регрессия



<https://nchu-datalab.github.io>

Векторная форма

- ▶ Пусть дан набор из n точек: $\{x_i, y_i\}_{i=1}^n$, где
 - $x_i = (x_{i1}, x_{i2}, \dots, x_{id})^T$ - вектор из d признаков объекта;
 - y_i - скалярная величина, которую хотим предсказать для объекта.

- ▶ Модель линейной регрессии:

$$\hat{y}_i = w_0 + \sum_{j=1}^d w_j x_{ij}$$

- w_j - веса модели;
 - \hat{y}_i - прогноз для объекта;
- ▶ Квадрат ошибки прогноза модели для объекта: $(\hat{y}_i - y_i)^2$

Матричная форма

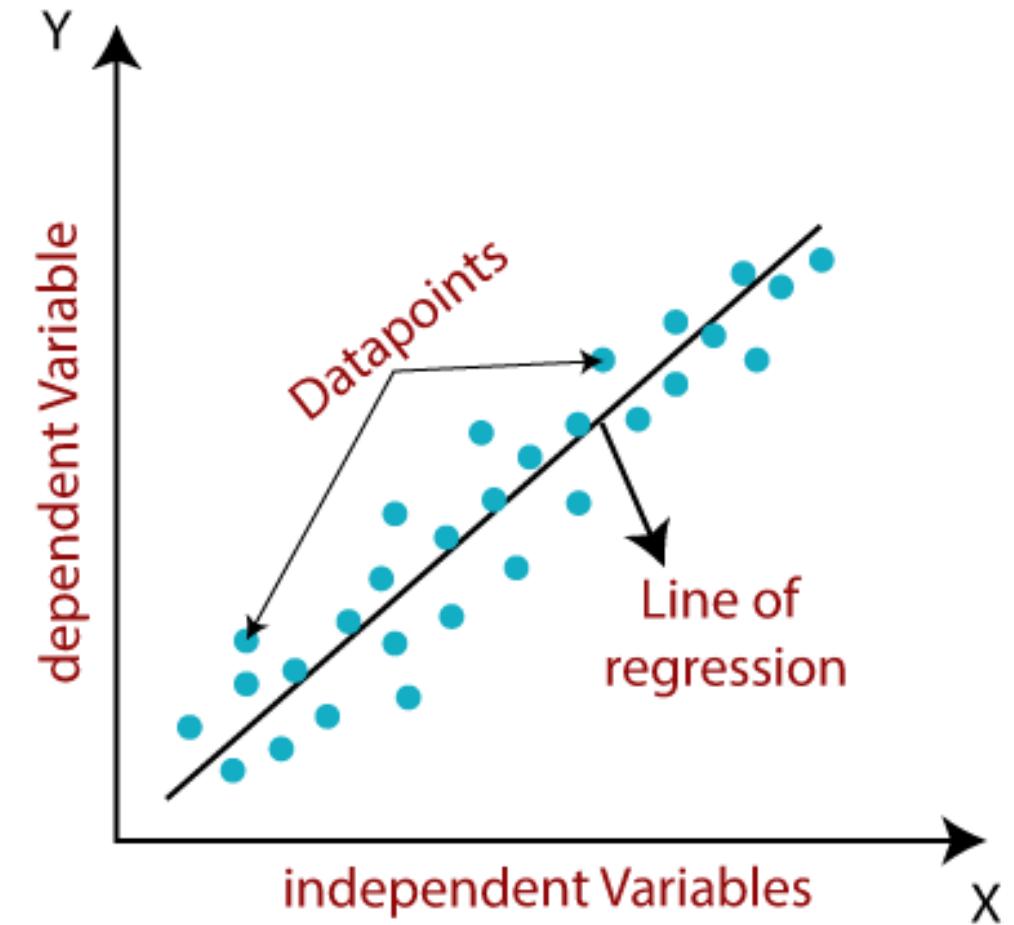
- ▶ Модель линейной регрессии:

$$\hat{y} = Xw$$

- $X = \begin{pmatrix} \mathbf{1} & x_{11} & \cdots & x_{1d} \\ \vdots & \ddots & & \vdots \\ \mathbf{1} & x_{n1} & \cdots & x_{nd} \end{pmatrix}$ - матрица признаков объектов $(n, d + 1)$;
 - $w = (w_0, w_1, \dots, w_d)^T$ - вектор $(d + 1)$ весов модели;
 - $\hat{y} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n)^T$ - вектор прогнозов модели для (n) объектов;
-
- ▶ Вектор квадратов ошибок прогнозов модели: $(\hat{y} - y)^2$

Задача

- ▶ Хотим, чтобы средняя квадратичная ошибка прогнозов $(\hat{y} - y)^2$ была минимальной
- ▶ **Как найти** оптимальные веса w модели?



Решение

- ▶ **Функция потерь (Loss function)** (скалярная и векторная формы):

$$L = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 = \frac{1}{n} (\hat{\mathbf{y}} - \mathbf{y})^T (\hat{\mathbf{y}} - \mathbf{y})$$

- ▶ Значение L – среднеквадратичная ошибка (**Mean Squared Error (MSE)**)
- ▶ Мы хотим минимизировать L :

$$L \rightarrow \min_w$$

Аналитическое решение

$$L = (\hat{y} - y)^T (\hat{y} - y) = (\mathbf{X}\mathbf{w} - y)^T (\mathbf{X}\mathbf{w} - y)$$

Чтобы найти минимум L , надо:

$$\frac{\partial L}{\partial \mathbf{w}} = 0$$

Тогда

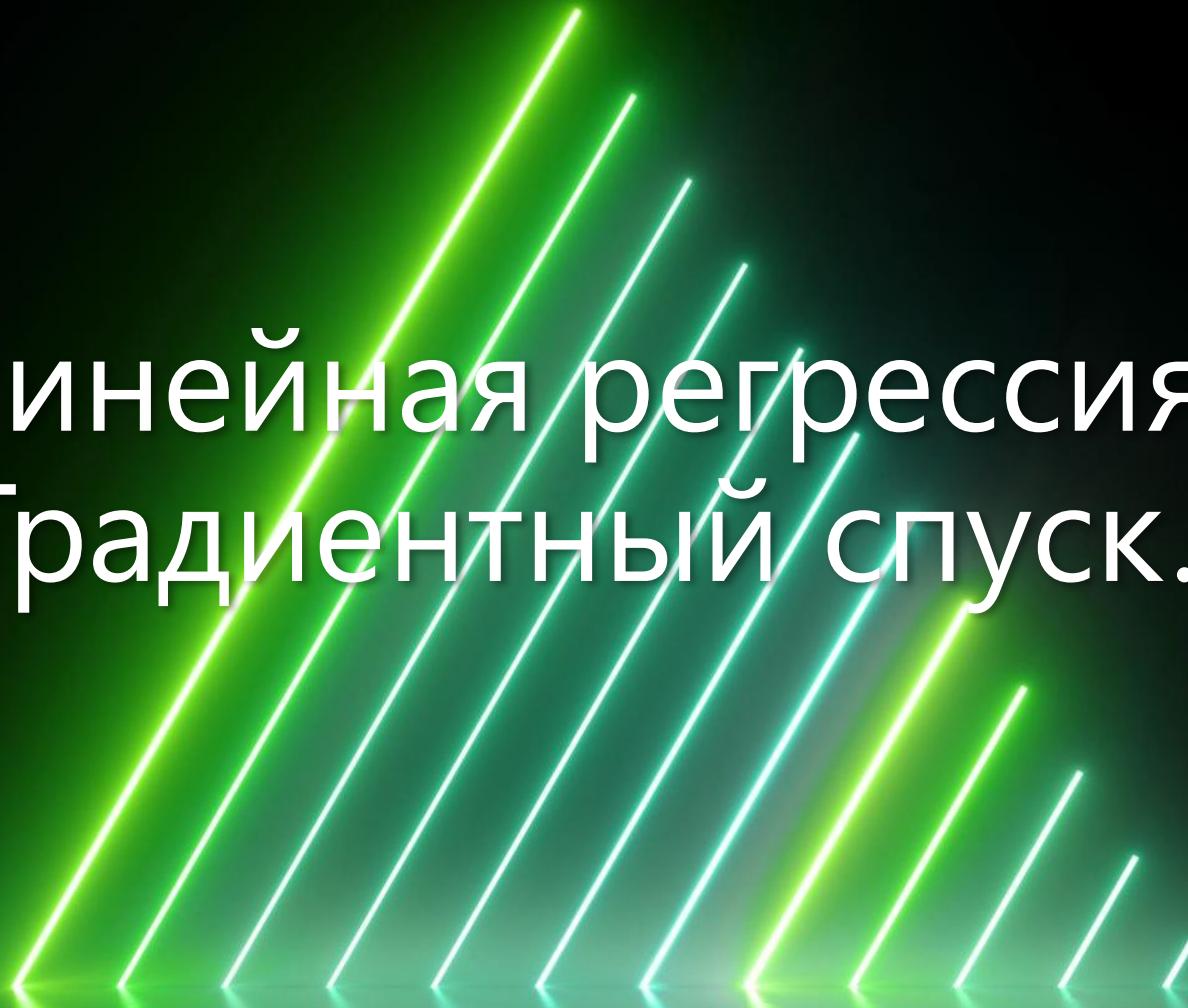
$$\frac{\partial L}{\partial \mathbf{w}} = 2\mathbf{X}^T(\mathbf{X}\mathbf{w} - y) = 2\mathbf{X}^T\mathbf{X}\mathbf{w} - 2\mathbf{X}^Ty = 0$$

Получаем оптимальные веса \mathbf{w} линейной регрессии:

$$\mathbf{w} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$$

Шпаргалка по дифференцированию матриц: http://nabatchikov.com/blog/view/matrix_der

Линейная регрессия.
Градиентный спуск.

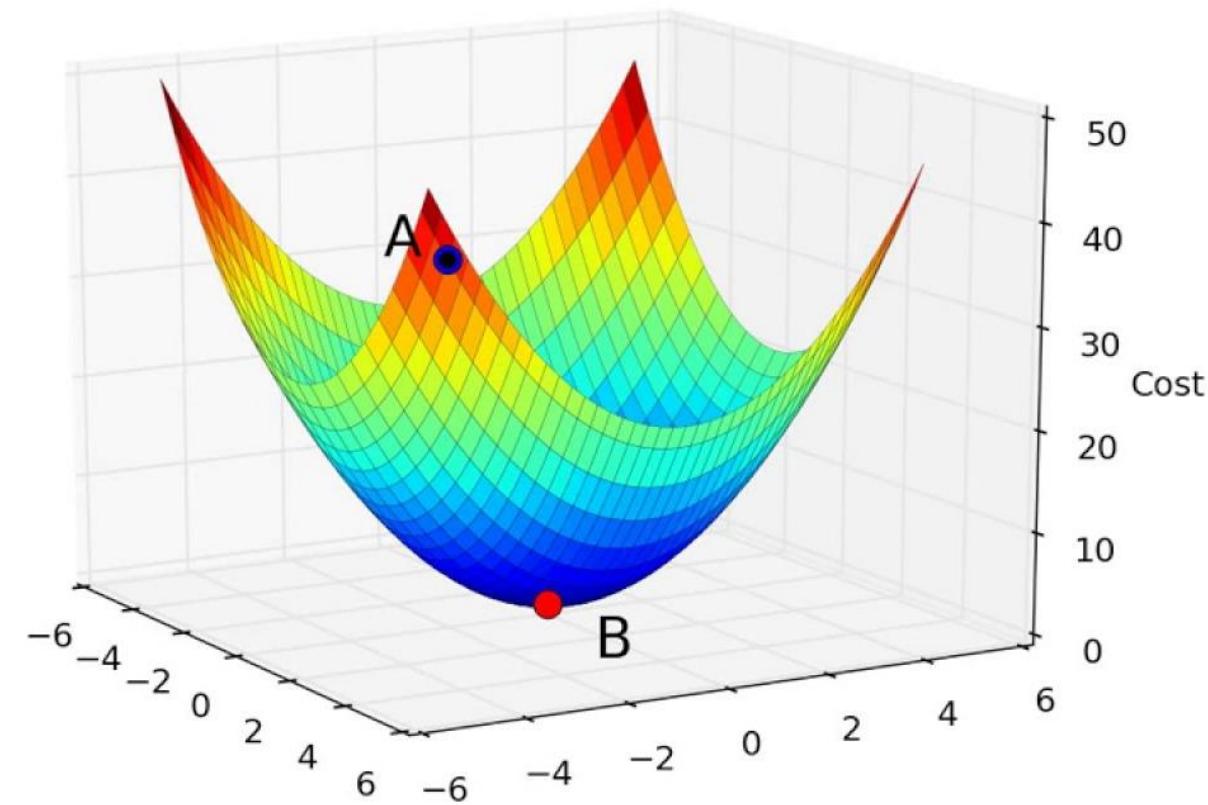


Задача

- ▶ Есть функция $L(w)$
- ▶ Хотим найти ее минимум:

$$L \rightarrow \min_w$$

- ▶ Мы умеем считать ее производную $\frac{\partial L}{\partial w}$
- ▶ Но не умеем (или не хотим) решать уравнение $\frac{\partial L}{\partial w} = 0$



Градиент функции

- ▶ Градиент функции (∇L) – вектор первых частных производных функции:

$$\nabla L(w_0, w_1, \dots, w_d) = \left(\frac{\partial L}{\partial w_0}, \frac{\partial L}{\partial w_1}, \dots, \frac{\partial L}{\partial w_d} \right)$$

- ▶ В векторной форме мы будем писать:

$$\nabla L(w) = \frac{\partial L(w)}{\partial w}$$

Градиентный спуск

- ▶ Есть функция $L(w)$, минимум которой хотим найти
- ▶ Пусть $w^{(0)}$ - начальный вектор параметров. Например, $w^{(0)} = 0$
- ▶ Тогда **градиентный спуск** состоит в повторении:

$$w^{(k+1)} = w^{(k)} - \eta \nabla L(w^{(k)})$$

- η – длина шага градиентного спуска (**learning rate**) (мы сами его задаем)
- k – номер итерации
- $\nabla L(w^{(k)})$ – градиент функции потерь на итерации k

Пример

- ▶ Модель линейной регрессии:

$$\hat{y} = Xw$$

- ▶ Функция потерь MSE:

$$L = \frac{1}{n} (\hat{y} - y)^T (\hat{y} - y) = \frac{1}{n} (Xw - y)^T (Xw - y)$$

- ▶ Градиент:

$$\nabla L(w) = \frac{2}{n} X^T (Xw - y)$$

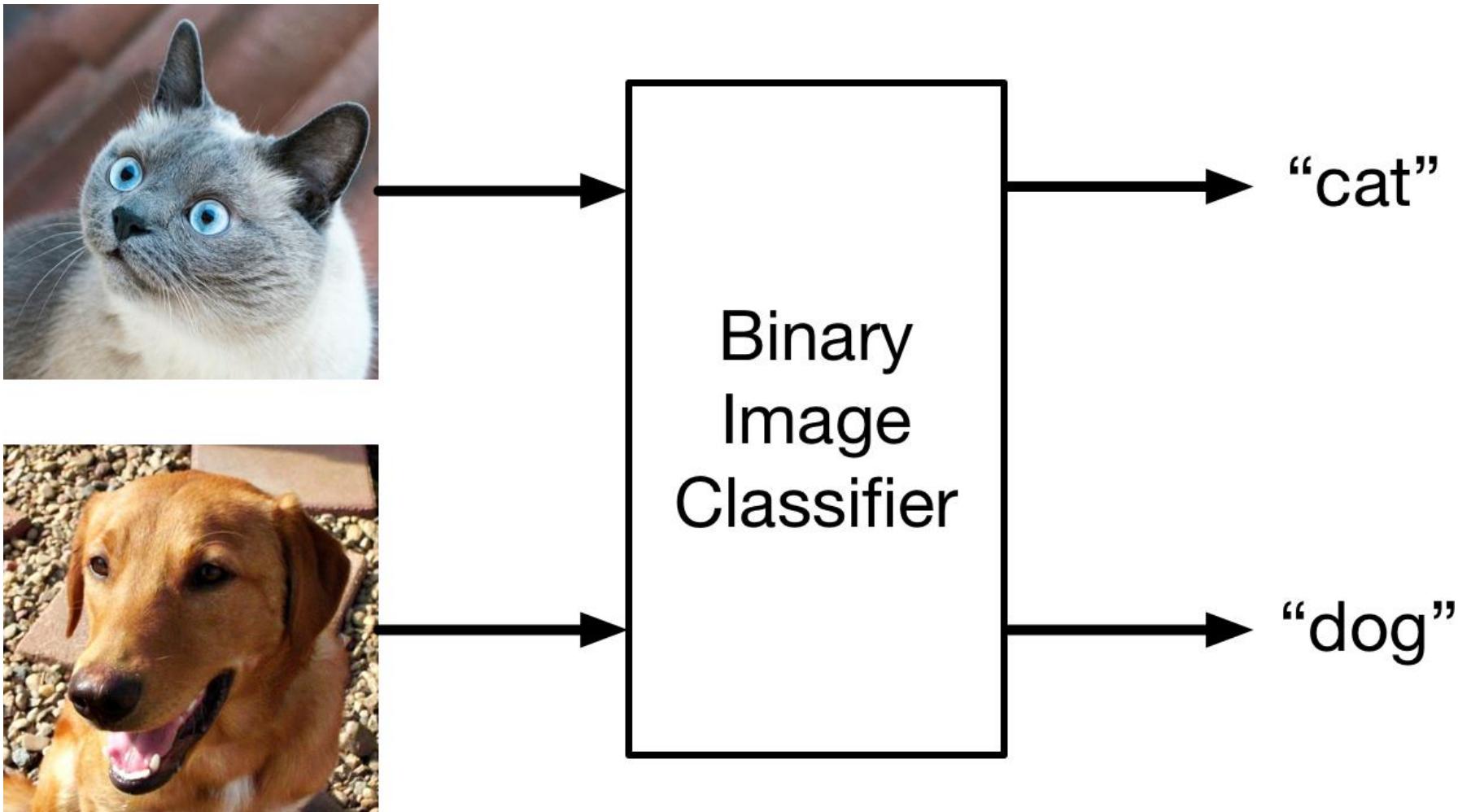
- ▶ Градиентный спуск:

$$w^{(k+1)} = w^{(k)} - \eta \nabla L(w^{(k)}) = w^{(k)} - \eta \frac{2}{n} X^T (Xw^{(k)} - y)$$

Классификация



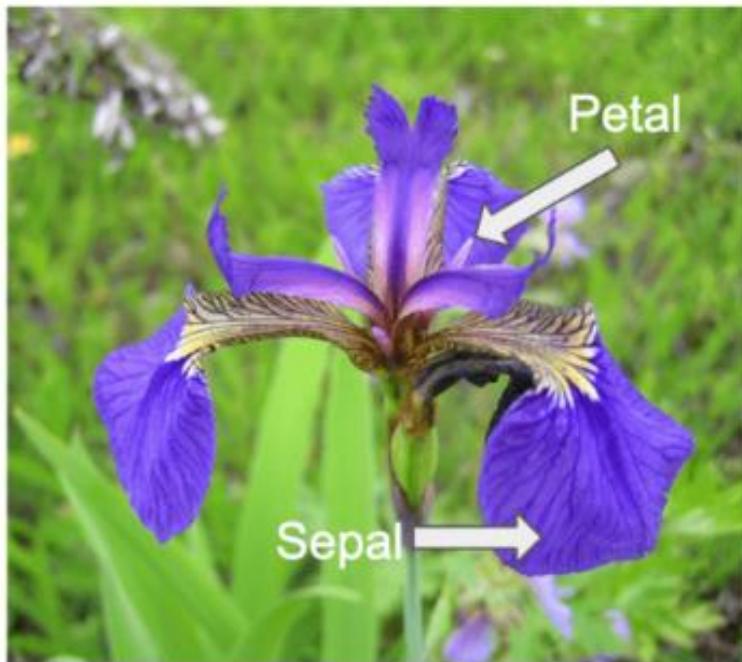
Классификация изображений



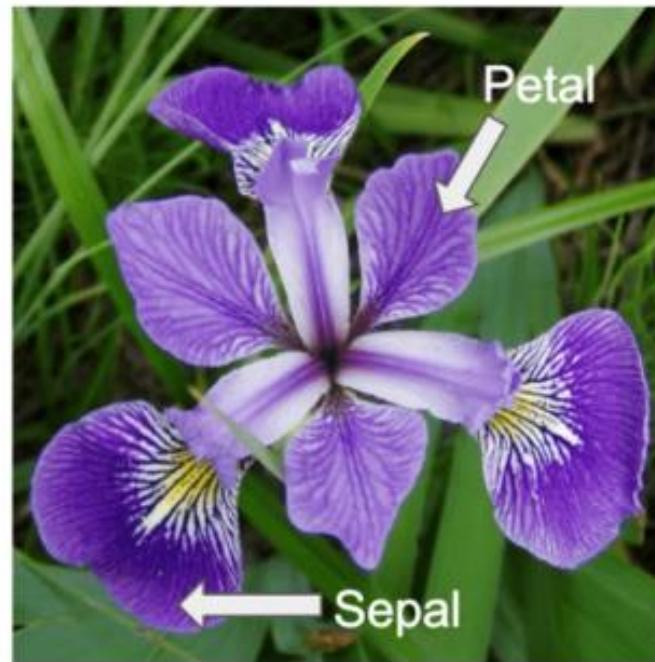
<https://www.raywenderlich.com/>

Классификация ирисов

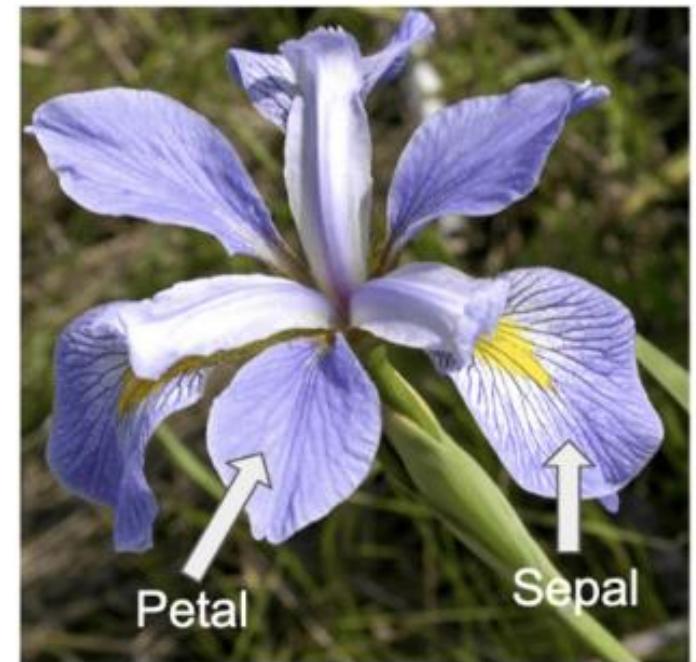
Iris setosa



Iris versicolor



Iris virginica

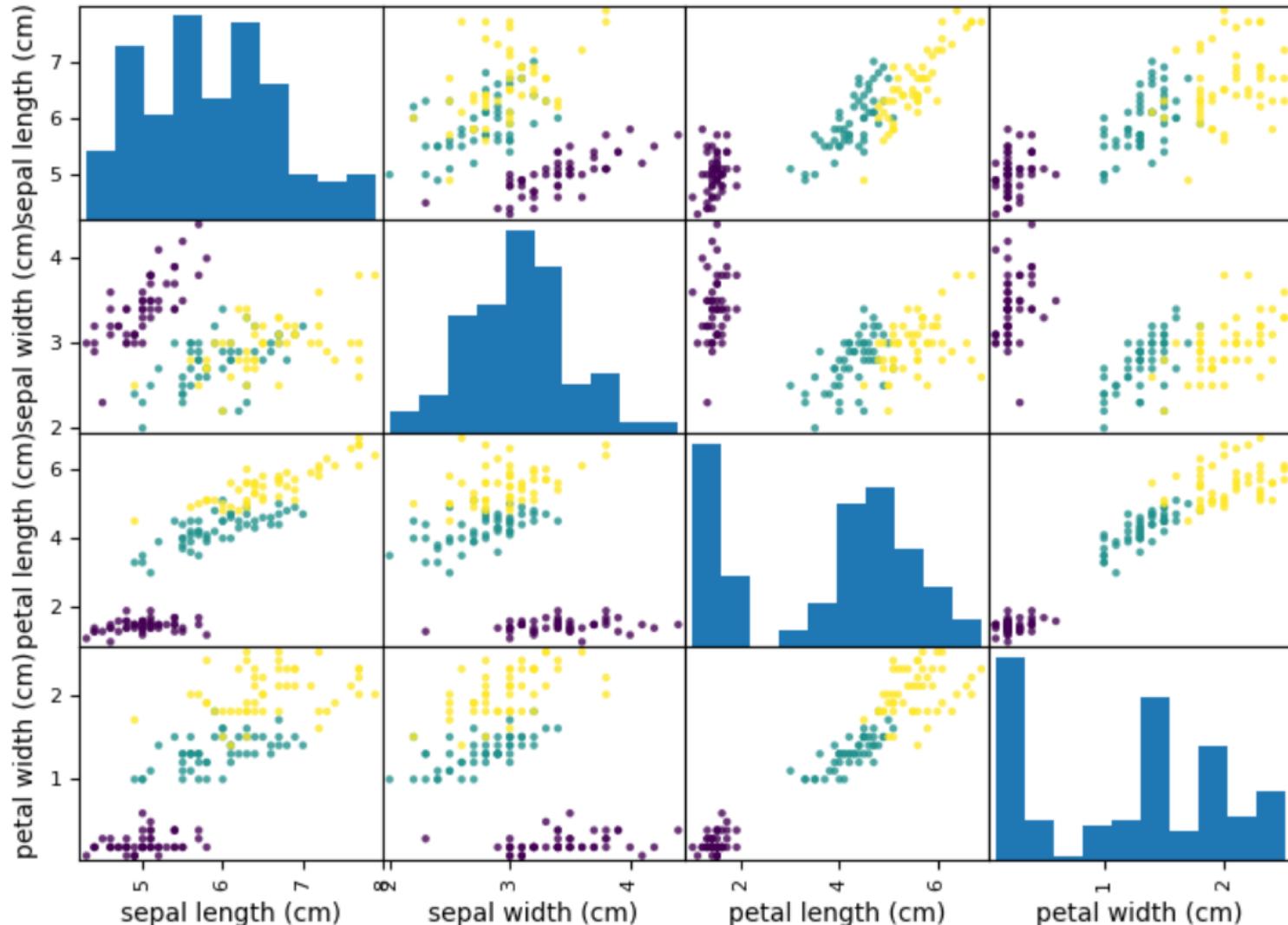


Классификация ирисов

Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
1	5.1	3.5	1.4	0.2	Iris-setosa
2	4.9	3	1.4	0.2	Iris-setosa
3	4.7	3.2	1.3	0.2	Iris-setosa
4	4.6	3.1	1.5	0.2	Iris-setosa
5	5	3.6	1.4	0.2	Iris-setosa
6	5.4	3.9	1.7	0.4	Iris-setosa
7	4.6	3.4	1.4	0.3	Iris-setosa
8	5	3.4	1.5	0.2	Iris-setosa
9	4.4	2.9	1.4	0.2	Iris-setosa
10	4.9	3.1	1.5	0.1	Iris-setosa
11	5.4	3.7	1.5	0.2	Iris-setosa
12	4.8	3.4	1.6	0.2	Iris-setosa
13	4.8	3	1.4	0.1	Iris-setosa
14	4.3	3	1.1	0.1	Iris-setosa
15	5.8	4	1.2	0.2	Iris-setosa
16	5.7	4.4	1.5	0.4	Iris-setosa
17	5.4	3.9	1.3	0.4	Iris-setosa
18	5.1	3.5	1.4	0.3	Iris-setosa
19	5.7	3.8	1.7	0.3	Iris-setosa



Классификация ирисов

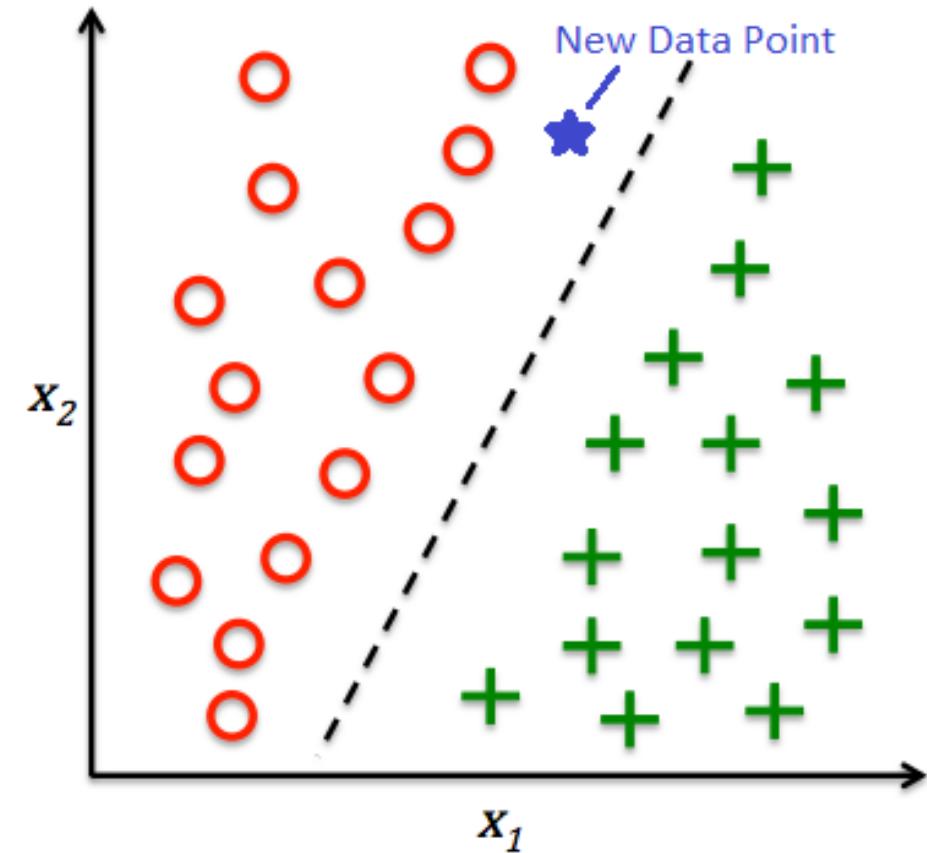


Логистическая регрессия



Задача

- ▶ Есть объекты двух классов
- ▶ Нужно разделить объекты по классам некоторой **гиперплоскостью**
- ▶ Эту гиперплоскость будем называть **линейным классификатором**



Матричная форма

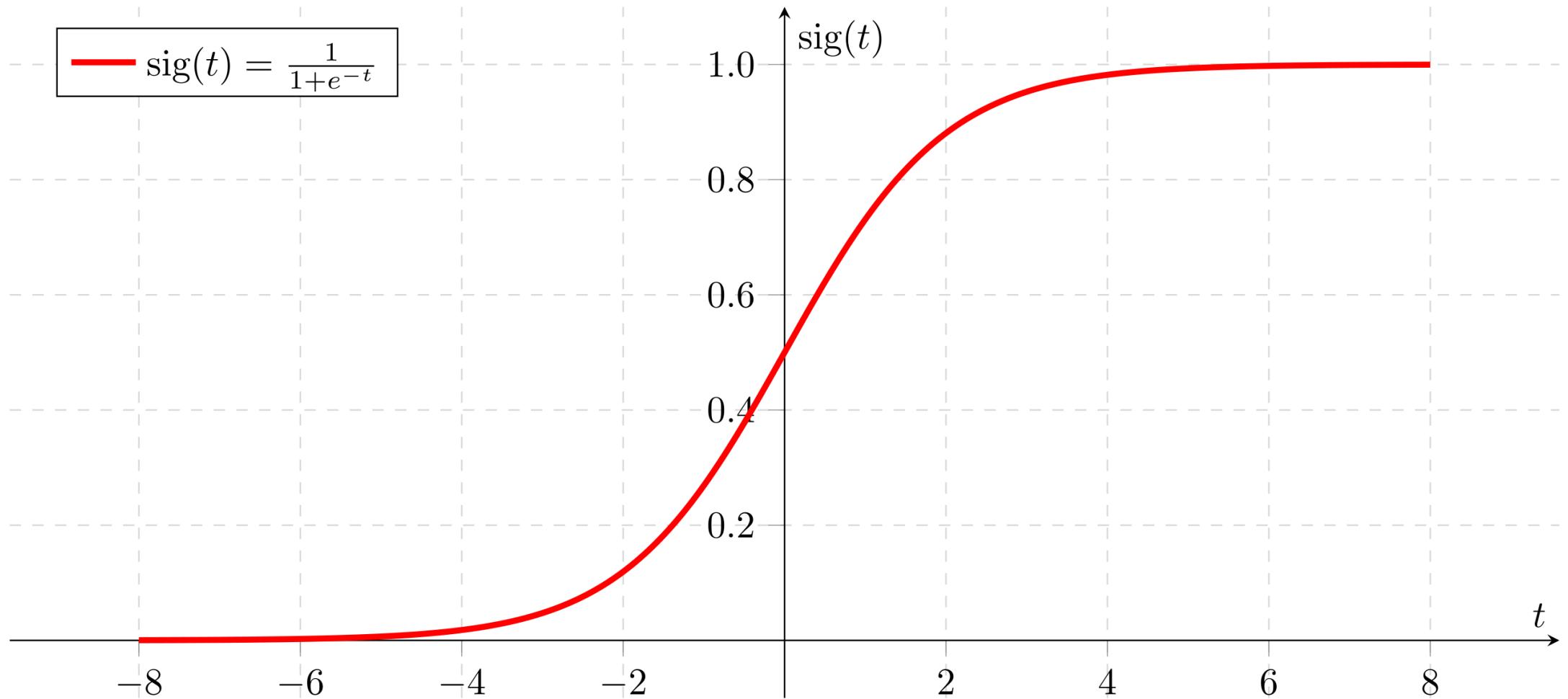
- ▶ Пусть дан набор из n точек: $\{x_i, y_i\}_{i=1}^n$, где
 - $x_i = (x_{i1}, x_{i2}, \dots, x_{id})^T$ – вектор из d признаков объекта;
 - $y_i = \{\mathbf{0}, \mathbf{1}\}$ – метка класса объекта.
- ▶ Модель логистической регрессии:

$$\hat{y}_i = \sigma(Xw)$$

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

- ▶ \hat{y}_i - **вероятность класса 1** для объекта;

Сигмоида



Функция потерь

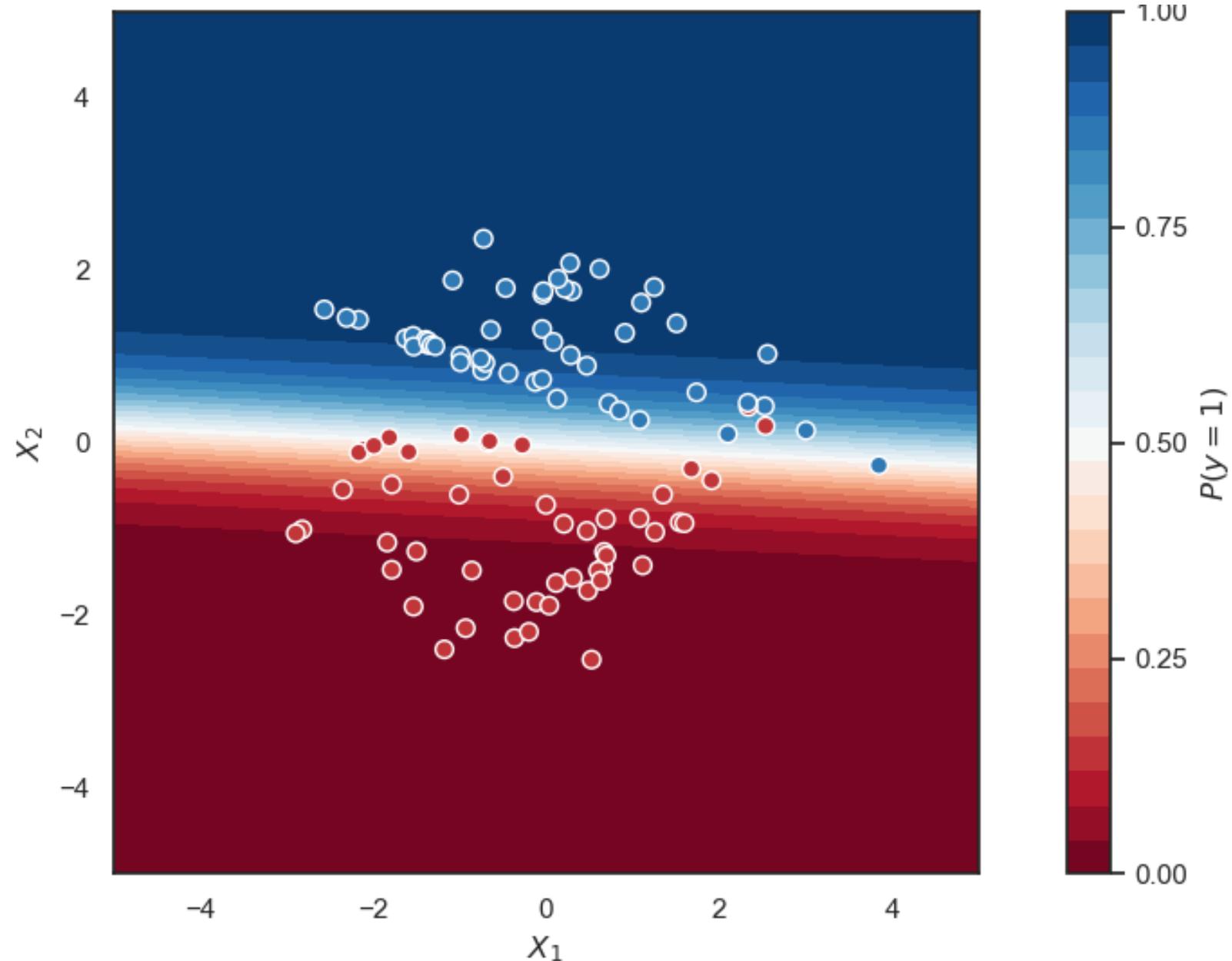
- ▶ Функция потерь для логистической регрессии (**log-loss**):

$$L = -\frac{1}{n} \sum_{i=1}^n (\textcolor{red}{y}_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i))$$

- ▶ Мы хотим минимизировать L :

$$L \rightarrow \min_w$$

Пример



Повтор

- ▶ Модель логистической регрессии:

$$\hat{y} = \sigma(Xw)$$

- ▶ Функция потерь log-loss:

$$L = -\frac{1}{n} \sum_{i=1}^n (y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i))$$

- ▶ Мы хотим минимизировать L :

$$L \rightarrow \min_w$$

- ▶ Градиентный спуск:

$$w^{(k+1)} = w^{(k)} - \eta \nabla L(w^{(k)})$$