

Машинное обучение

Илья Езепов

18. ... Bxe7



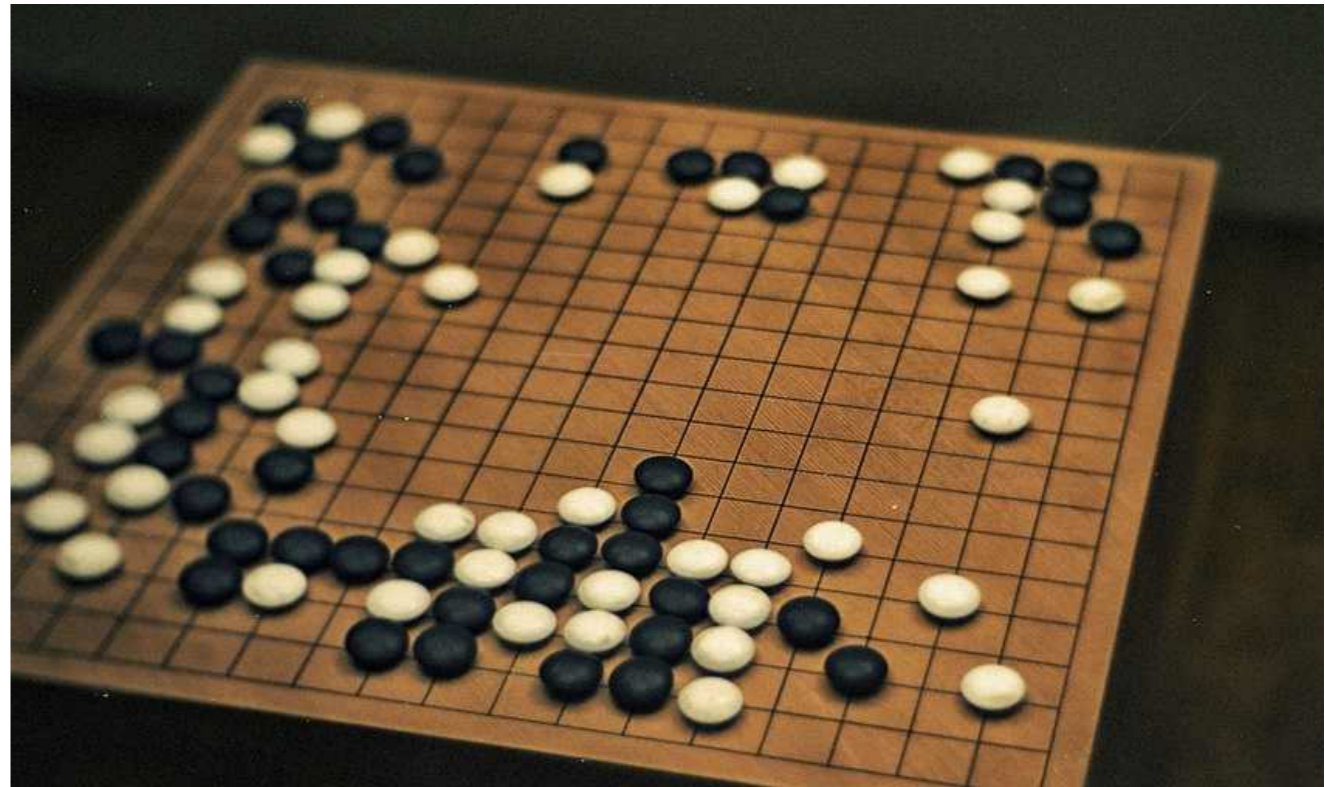
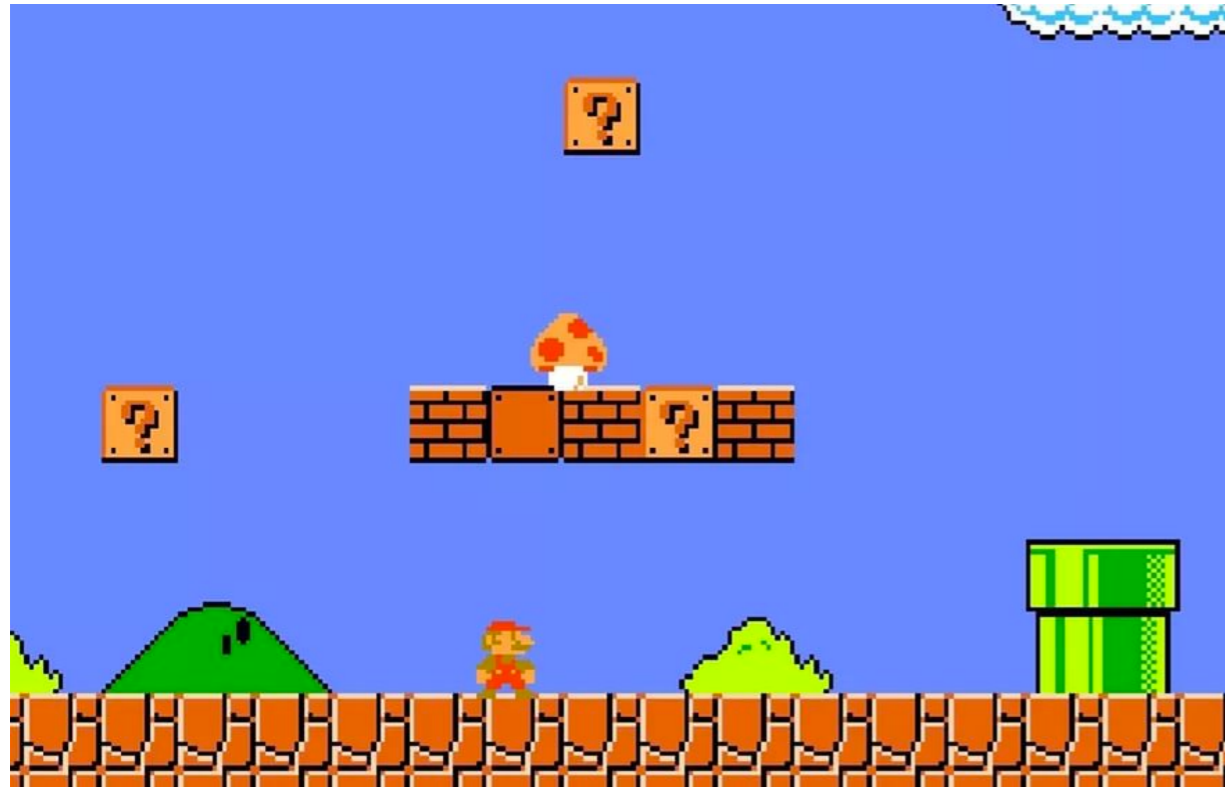
Ход белых

19.c4









Машинное обучение –
алгоритмы, способные обучаться

amazon.com[®]

The Amazon logo, a curved orange arrow pointing from the letter 'a' to the letter 'z', is positioned below the text 'amazon.com'.



Стократное увеличение продаж
с автоматических рекомендаций

**Объект, прецедент,
instance**

Ответ, target, label



Кошка



Собака




Кошка



?

Мы говорим хотя бы про сотни объектов, а порой про миллиарды

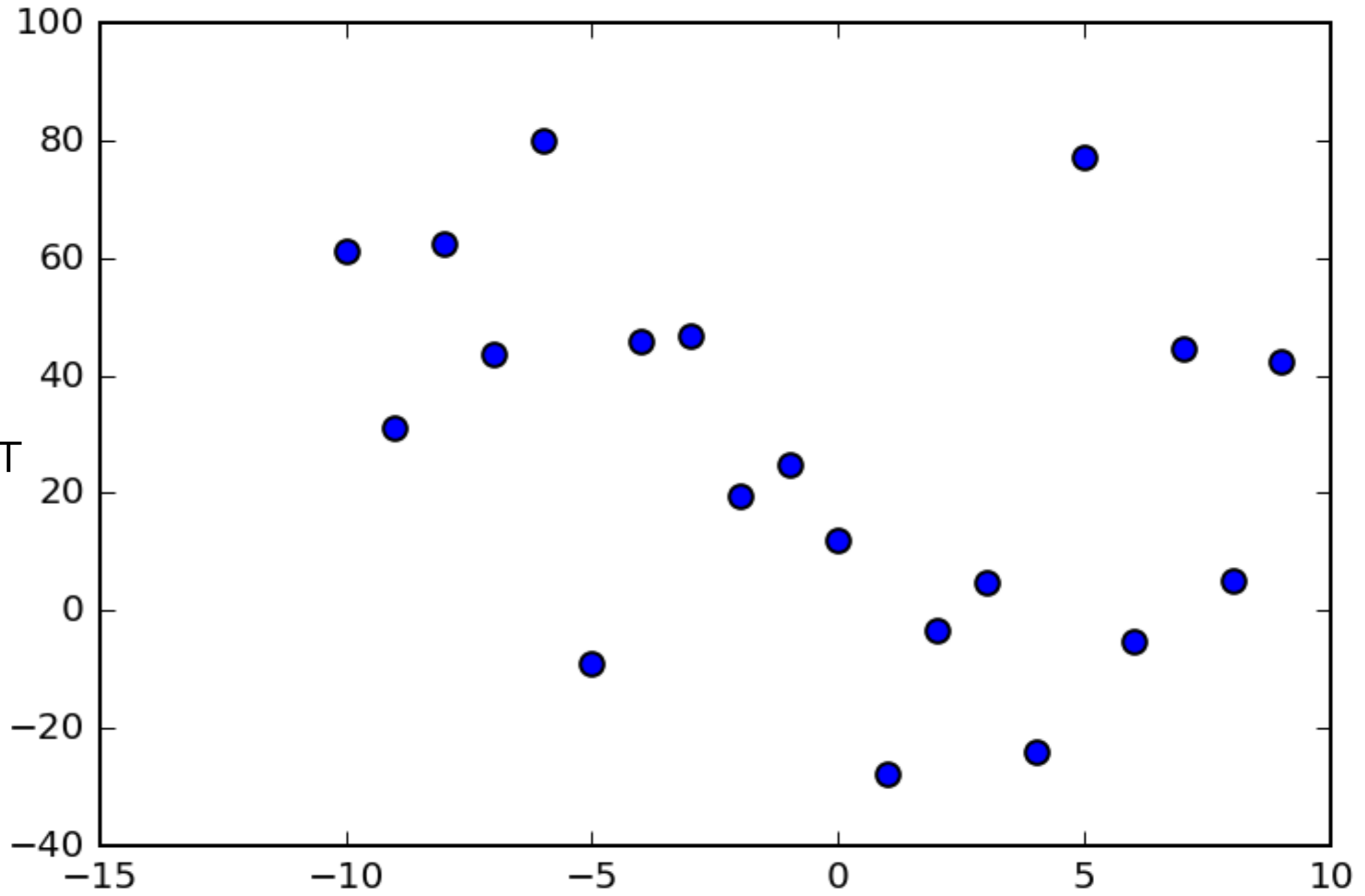
Признаки, факторы, features, фичи

	Масса	Количество лап	Цвет	Ответ
	3.7	4	Серый	Кошка
	24.1	4	Золотистый	Собака
	4.2	4	Белый	Кошка
	6.3	4	Коричневый	?

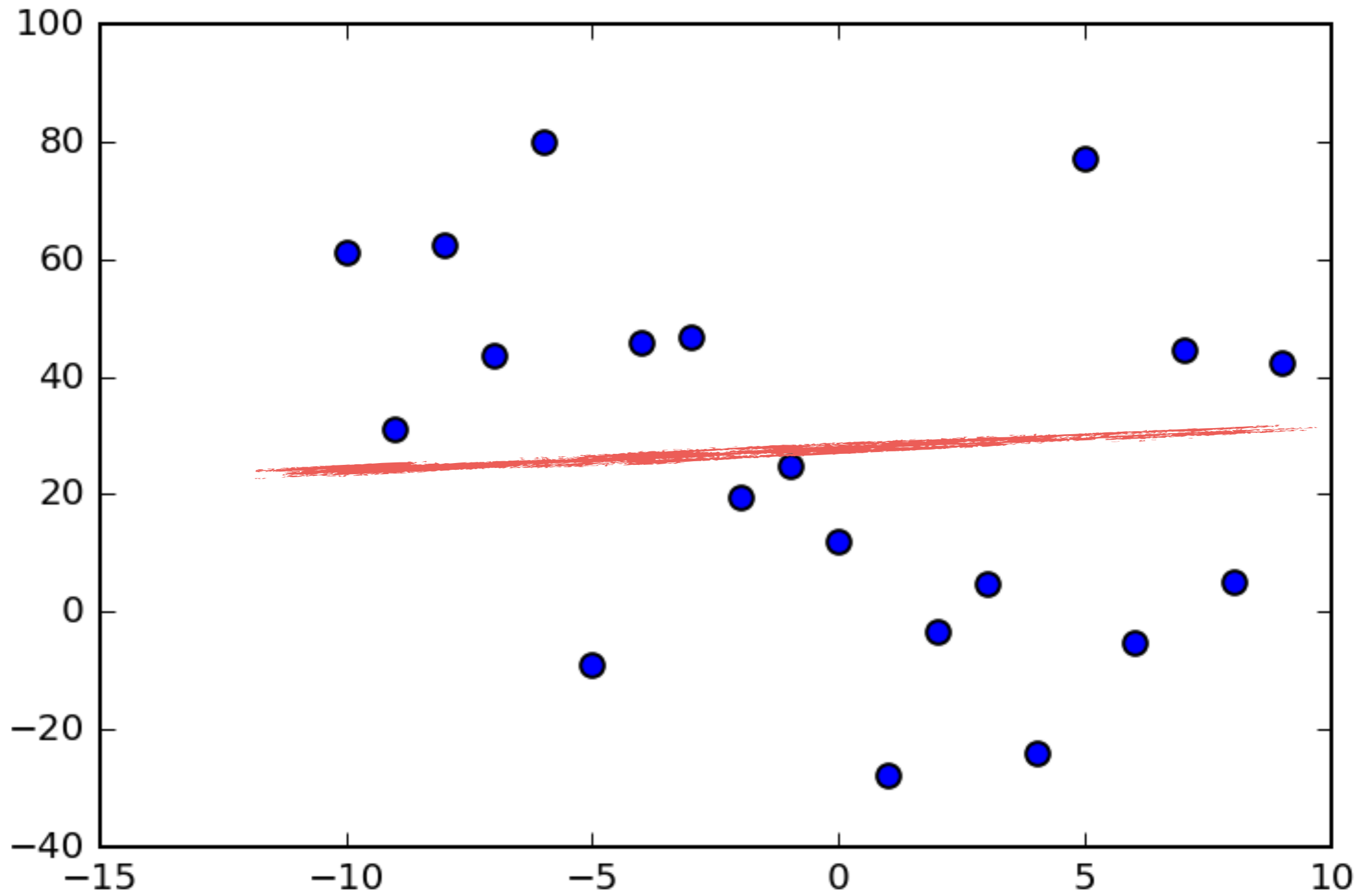
Признаки и ответы могут быть
самыми разными:

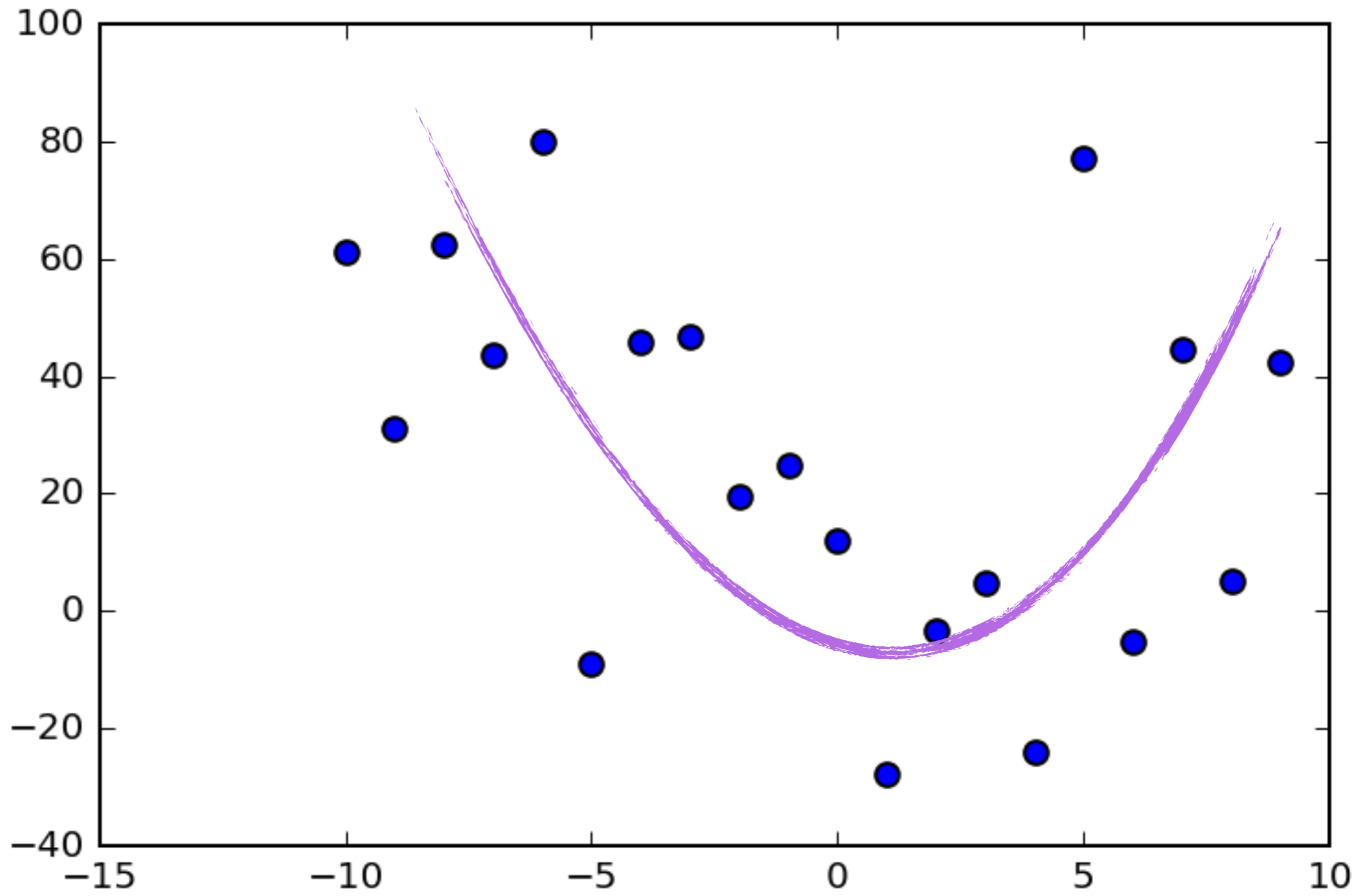
числа, категории, текст, картинки, ...

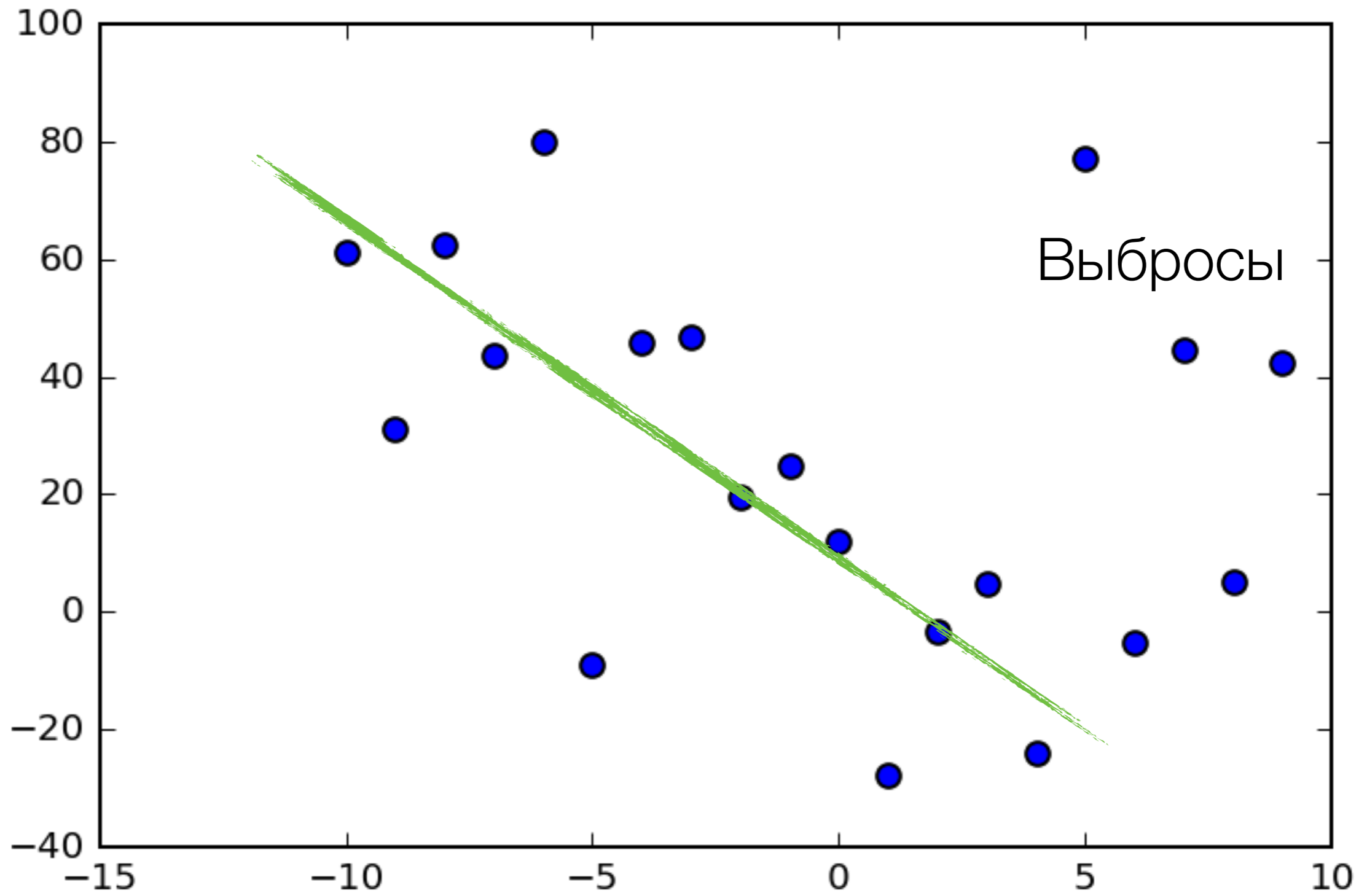
Ответ

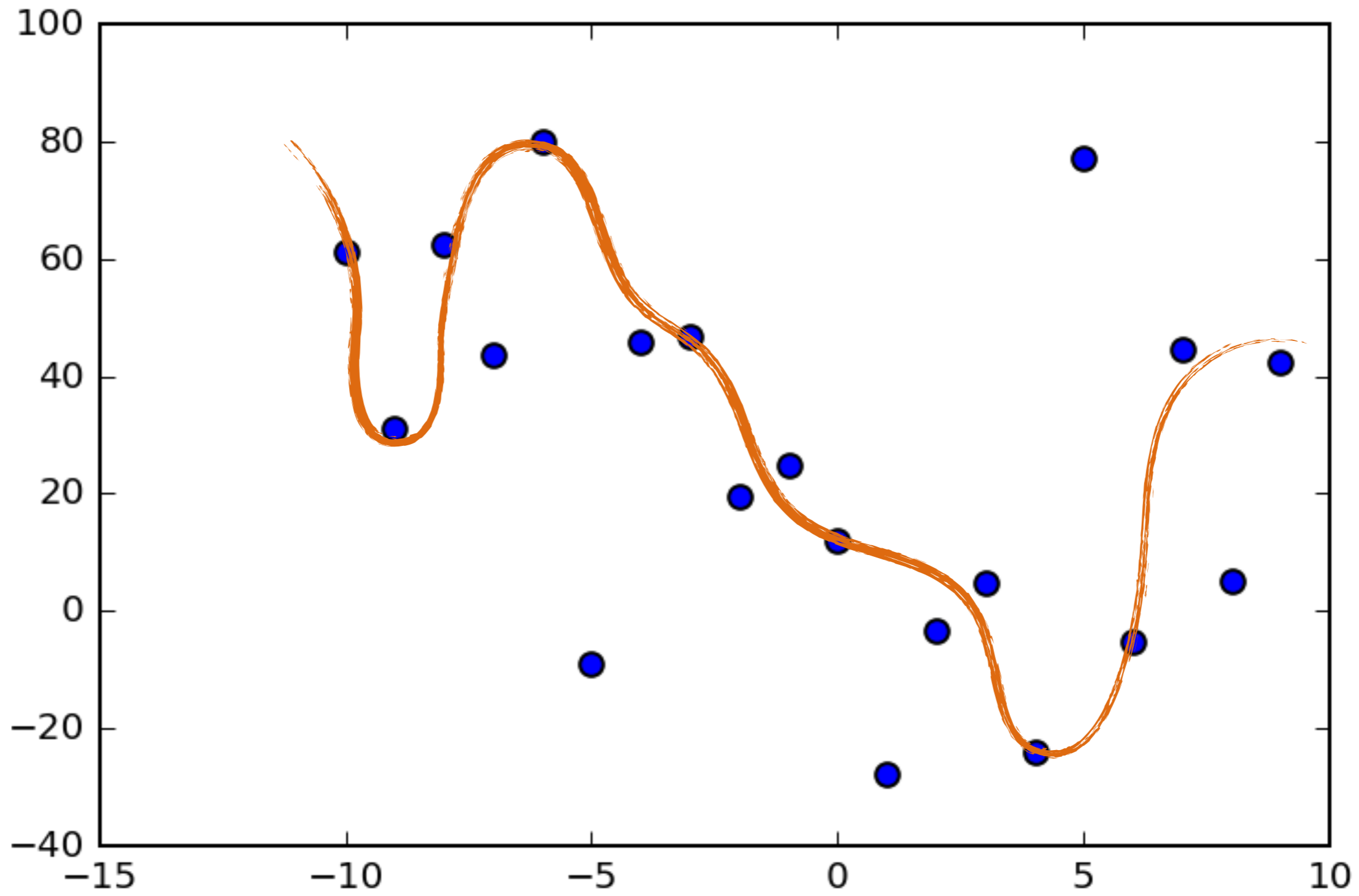


Признак

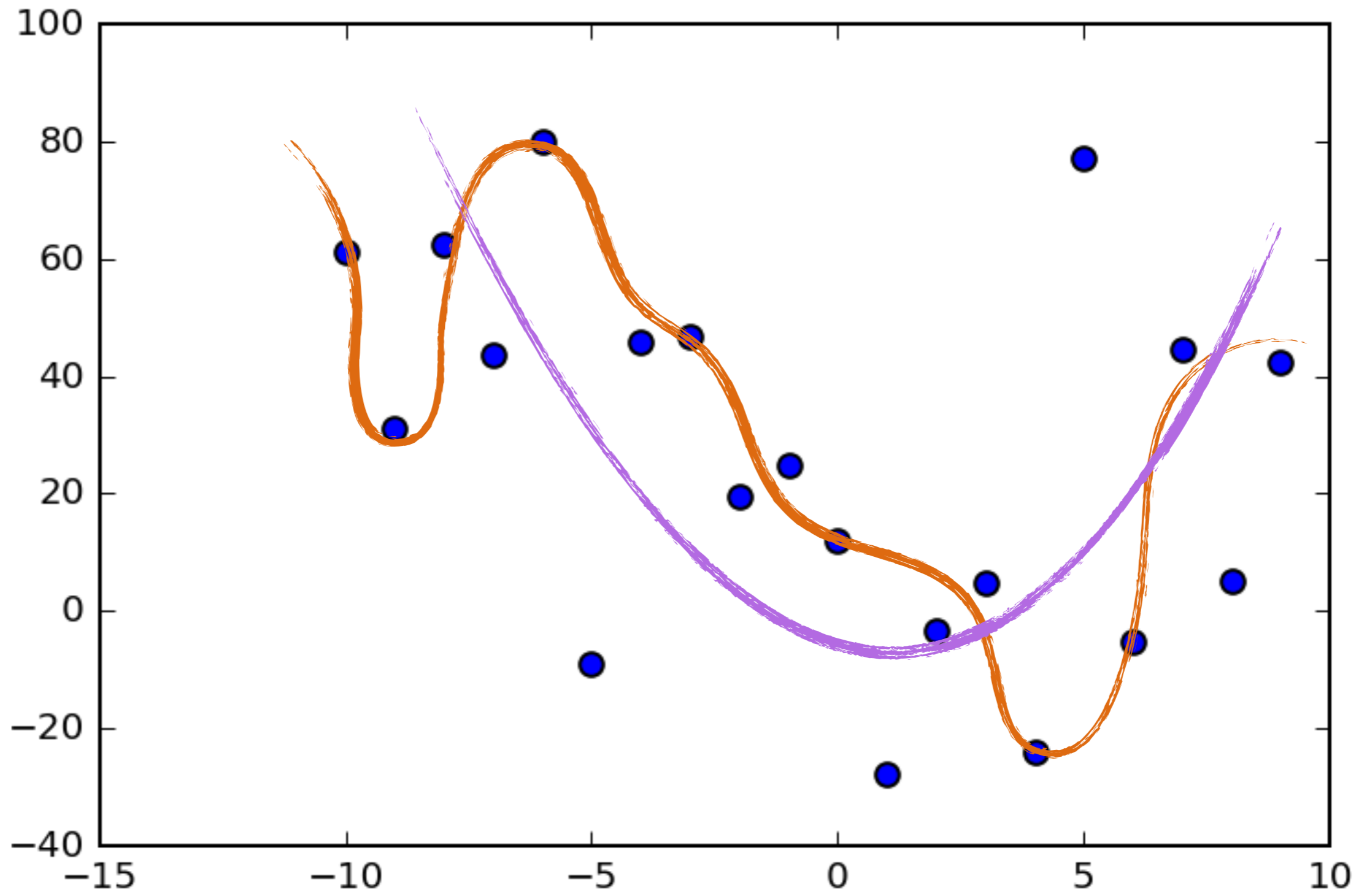








Провести кривую через точки
бывает сложно



Мы любим алгоритмы за
способность обобщать

Машинное обучение

Обучение с учителем

Обучение без учителя

Обучение с подкреплением

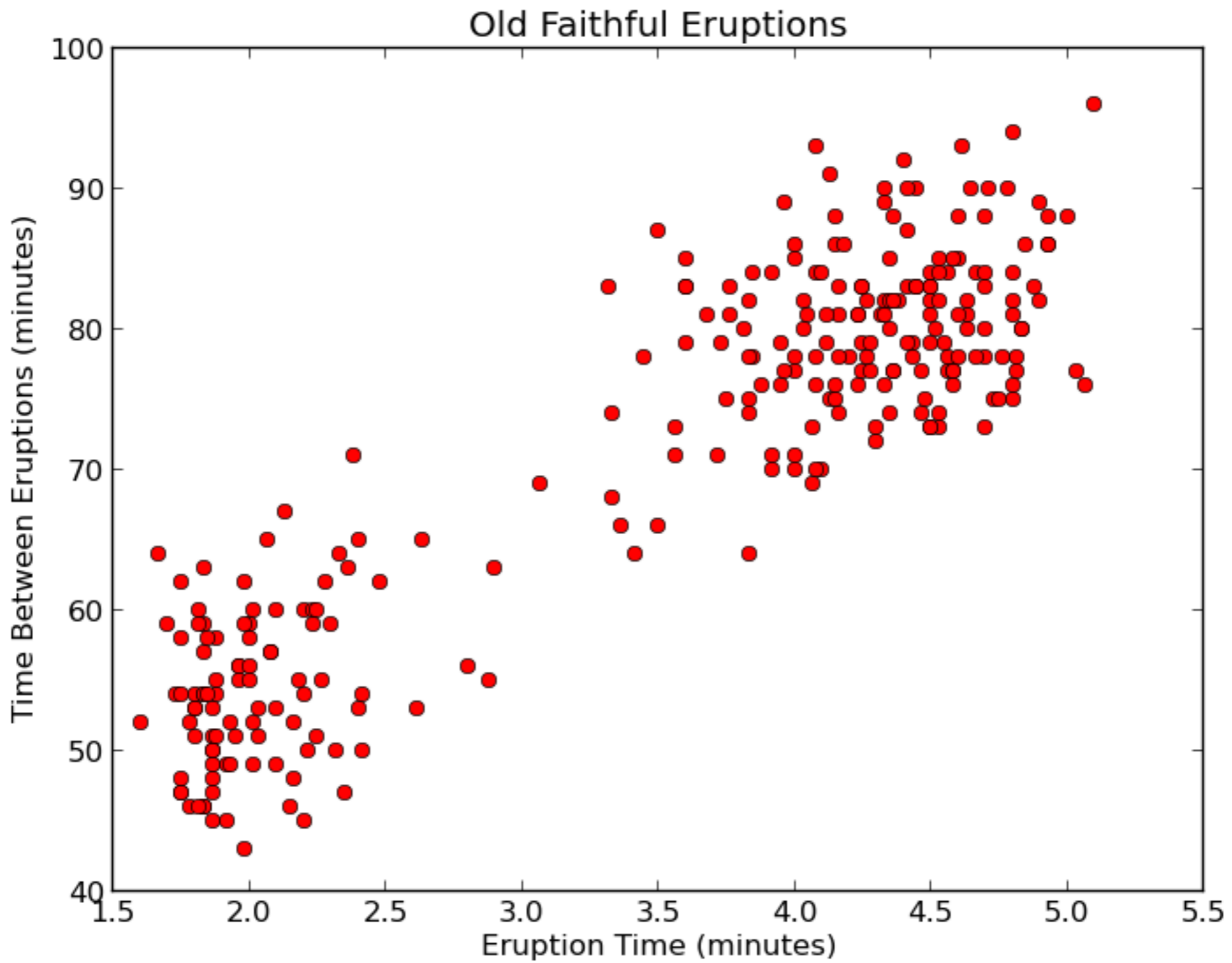
Машинное обучение

Обучение с
учителем

Обучение без
учителя

Обучение с
подкреплением





word2vec

- Алгоритм читает неразмеченный текст и выучивает вектора для каждого слова

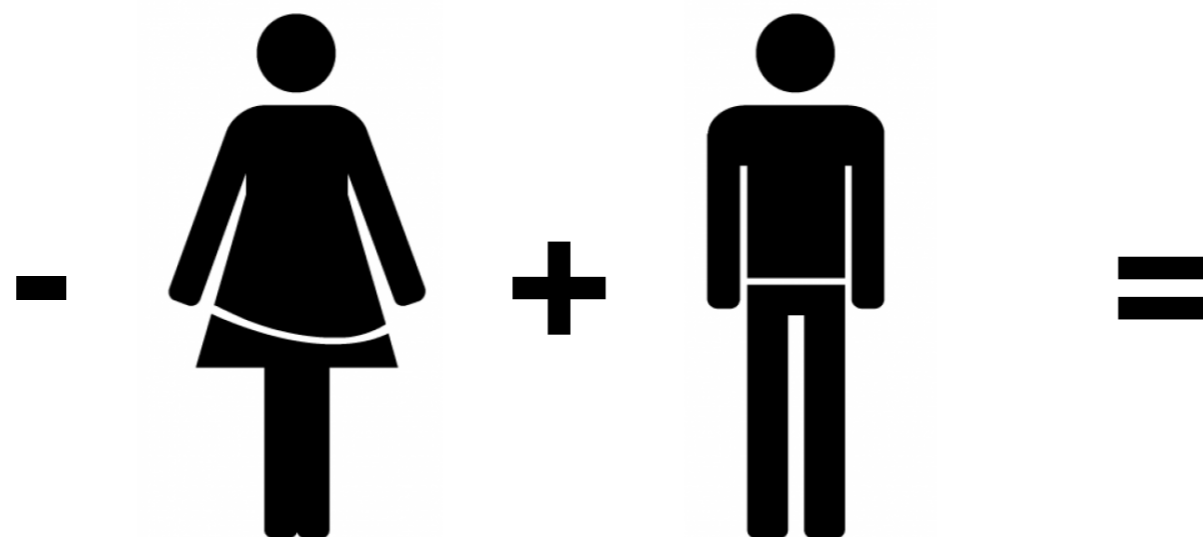
word2vec

- Похожие слова на *“кофе”*:
 - *“эспрессо”*
 - *“капучино”*
 - *“кофейный”*
 - *“чай”*
 - *“эспресо”*

word2vec



“королева”



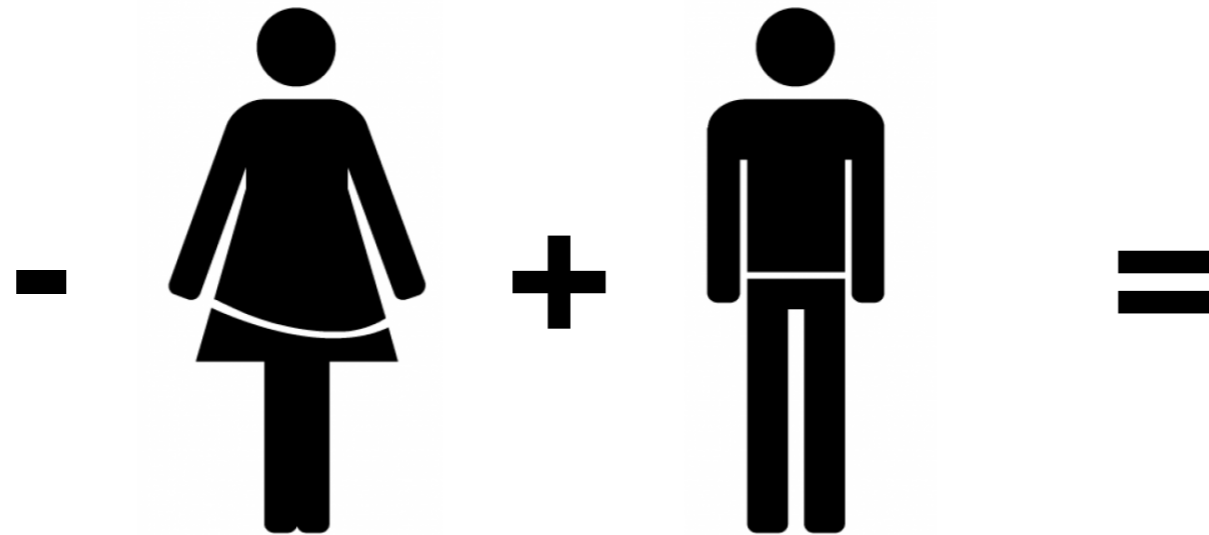
“женщина”

“мужчина”

word2vec



“королева”



“женщина”

“мужчина”



“король”

Машинное обучение

Обучение с
учителем

Обучение без
учителя

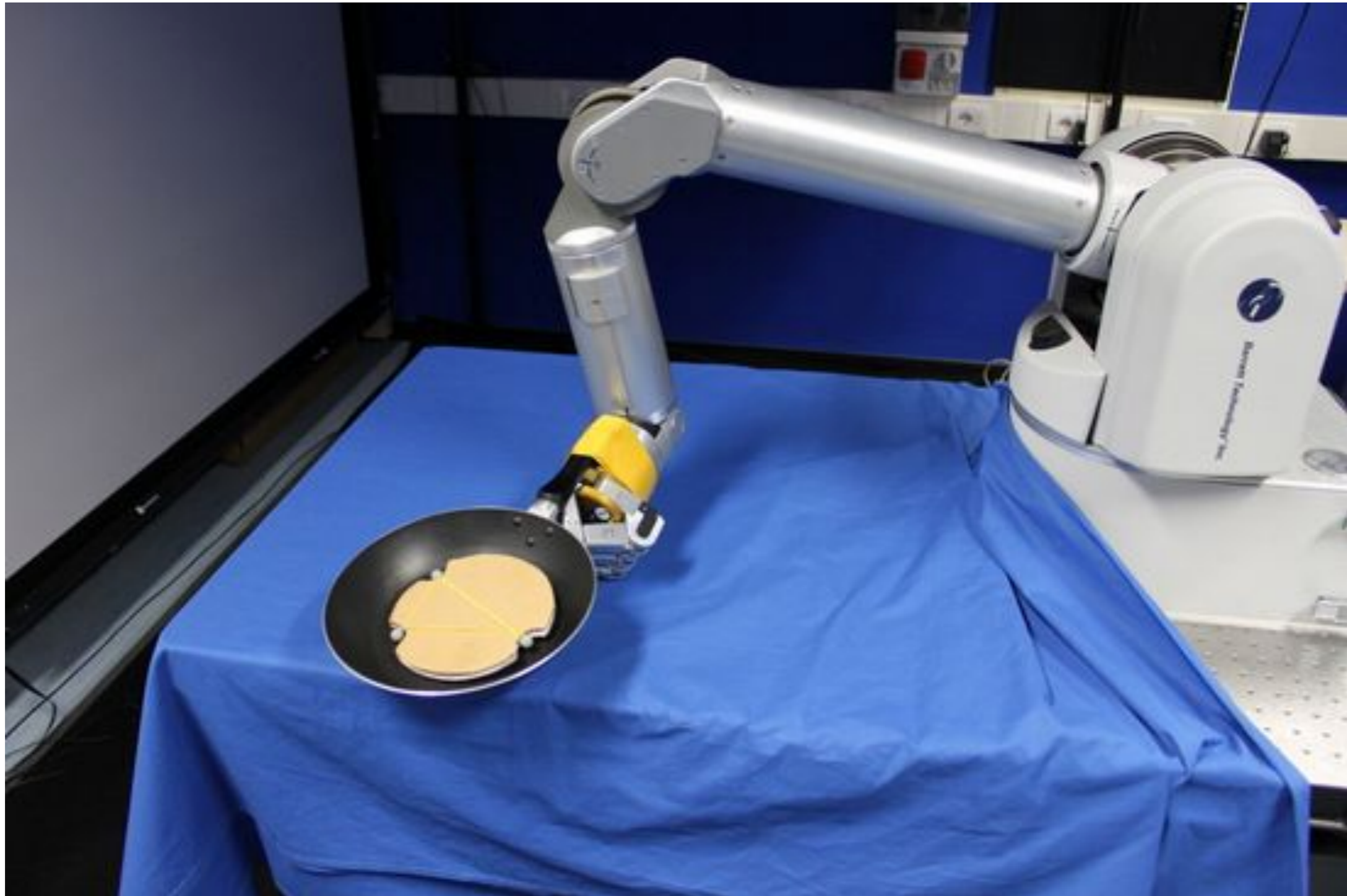
Обучение с
подкреплением

AlphaGo

- Сначала обучение с учителем:
пытаемся отгадать ходы мастеров
- Потом обучение с подкреплением:
разные версии алгоритма играют сами с собой,
становясь всё сильнее



Обратная связь



Машинное обучение
крайне многообразно

Анализ данных как спорт

- Идея: В 1998 Роб МакЭван обещает 500.000\$ за анализ данных для выбора места для золотого прииска.

Анализ данных как спорт

- Идея: В 1998 Роб МакЭван обещает 500.000\$ за анализ данных для выбора места для золотого прииска.

Зарабатывает \$3 млрд через год.

- Kaggle - главная мировая площадка спортивного анализа данных.

11 active competitions

Sort By

Prize

Active

All

Entered

Hosted

All Categories



Data Science Bowl 2017

Can you improve lung cancer detection?

Featured · 21 days to go

\$1,000,000

1,607 teams



The Nature Conservancy Fisheries Monitoring

Can you detect and classify species of fish?

Featured · 21 days to go

\$150,000

2,069 teams



Intel & MobileODT Cervical Cancer Screening

Which cancer treatment will be most effective?

Featured · 3 months to go

\$100,000

151 teams



Google Cloud Platform

Google Cloud & YouTube-8M Video Understanding Challenge

Can you produce the best video tag predictions?

Featured · 2 months to go

\$100,000

331 teams

<https://www.kaggle.com/competitions>

Что почитать на досуге

В этой книге много интересных примеров того, как сложнейшие технологии Big Data — методы анализа огромных объемов данных — применяются для решения важных задач из нашей повседневной жизни.

Сергей Мацоцкий, председатель правления компании IBS

ВИКТОР МАЙЕР-ШЕНБЕРГЕР | КЕННЕТ КУКЬЕР

BIG

БОЛЬШИЕ ДАННЫЕ

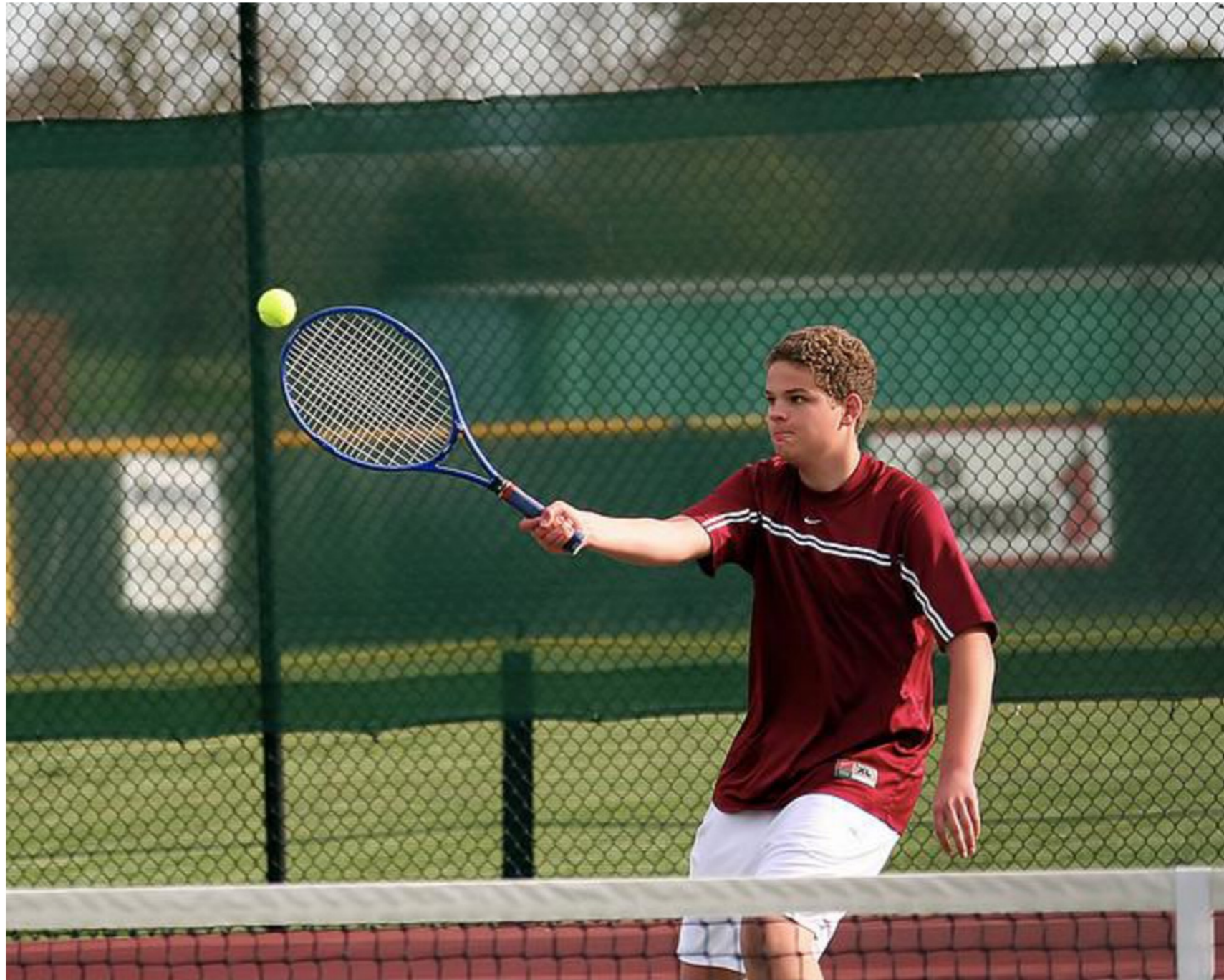
DATA

РЕВОЛЮЦИЯ, КОТОРАЯ
ИЗМЕНИТ ТО, КАК МЫ ЖИВЕМ,
РАБОТАЕМ И МЫСЛИМ

- Andrew Ng, Coursera
- Школа Анализа Данных, Яндекс

“Я скорее менеджер”

“Я скорее аналитик”



a man is playing tennis on a tennis court



a man sitting on a bus with a dog



a bunch of bananas are hanging from a ceiling