

Искусственный Интеллект

Лекция 1: Понятия Искусственного интеллекта и Машинного обучения

Мартынюк Полина Антоновна

telegram: @PAMartynyuk

email: pa-martynyuk@yandex.ru



Дисциплина ИИ

3 Модуля:

- I. Основы машинного обучения
РК 1: 18...30 баллов
- II. Методы подготовки данных и оценки качества моделей машинного обучения
РК2: 21...35 баллов
ЛР1, ЛР2
- III. Продвинутое использование алгоритмов машинного обучения
РК3: 21...35 баллов
ЛР3, ЛР4

Σ: 60...100 баллов
Зачет: ≥ 60 баллов + сданы ЛР

Время и место занятий:

ЗН пятница, 15:40 – 19:00, ауд. 805



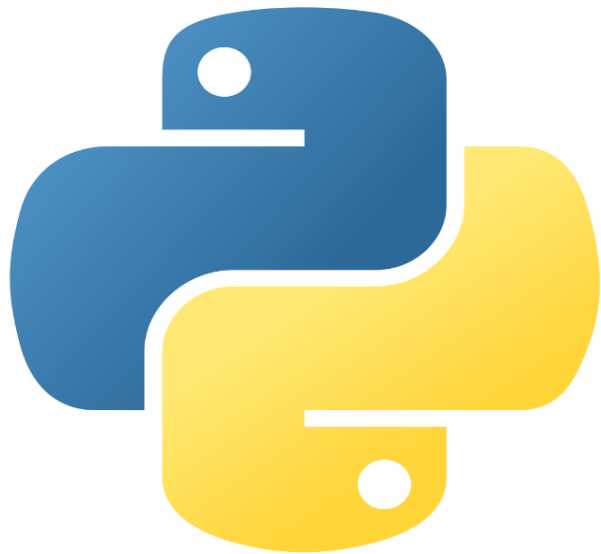
Группа для связи в telegram

Дисциплина ИИ

Темы дисциплины:

- Понятие искусственного интеллекта (ИИ). Основы машинного обучения (МО).
- Простейшие алгоритмы машинного обучения. Метод kNN.
- Обучение с учителем. Решение задачи регрессии. Перцептрон. Линейная регрессия.
- Градиентный спуск. Функция потерь.
- Обучение с учителем. Решение задачи классификации. Логистическая регрессия.
- Обучение без учителя. Кластеризация, сжатие данных.
- Способы предобработки исходных данных.
- Оценка модели МО. Переобучение моделей МО. Регуляризация.
- Ансамбли моделей МО. Деревья решений.
- Рекомендательные системы.
- *Основы нейросетевых архитектур.*

Инструменты



Python



Jupyter Notebook



Google Colab

Понятие ИИ

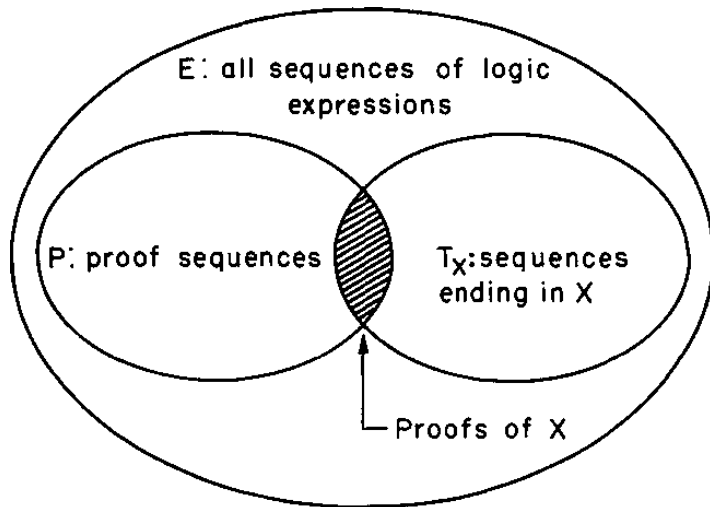


Что понимается под термином
«Искусственный Интеллект»?
(ИИ, Artificial intelligence, AI)



Первое определение ИИ

«ИИ - машины,
которым присуще
разумное
поведение»



Джон Маккарти — чудаковатый учёный-компьютерщик из Стэнфорда — ввёл термин «Искусственный интеллект» в **1956 году**.

В этот исторический год вместе с горсткой других учёных-математиков он провёл летний семинар на тему ИИ в Дартмутском колледже. Учёные с этой конференции в том же году создали первый ИИ в мире — компьютерную программу Logic Theorist. Она могла доказывать определенные математические теоремы. Так началась история искусственного интеллекта.

Logic Theorist мог доказывать теоремы символической логики и решать задачи логики высказываний и предикатов. Это удалось сделать путем поиска большого количества возможных решений и выбора того, которое лучше всего соответствовало ограничениям задачи.

Факты о программе:

- Logic Theorist создал лучшие доказательства, чем люди-математики.
- Его создание породило такие области исследований, как искусственный интеллект и эвристическое программирование.
- Его изобретение потребовало создания языка обработки информации (IPL).
- Превзошел математиков Альфреда Норта Уайтхеда и Бертрانا Рассела, поскольку предложил лучшие и более подробные доказательства их теорем.

Переосмысление определения ИИ

- В начале 1980-х гг. ученые в области теории вычислений Барр и Файгенбаум предложили следующее определение искусственного интеллекта (ИИ):

Искусственный интеллект — это область информатики, которая занимается разработкой интеллектуальных компьютерных систем, то есть систем, обладающих возможностями, которые мы традиционно связываем с человеческим разумом, — понимание языка, обучение, способность рассуждать, решать проблемы и т. д.

- Позже к ИИ стали относить ряд алгоритмов и программных систем, отличительным свойством которых является то, что они **могут решать некоторые задачи так, как это делал бы размышляющий над их решением человек.**

Понятие ИИ

- ИИ – область компьютерных наук о поиске неизвестных зависимостей (в мире = в данных)

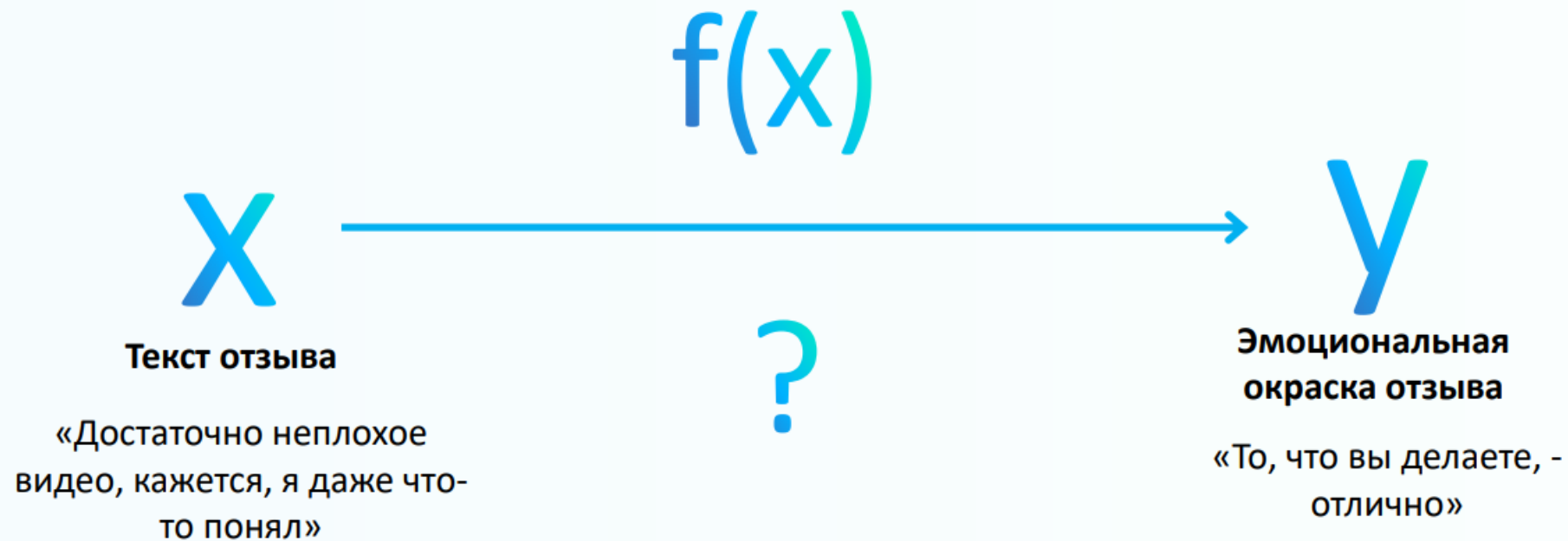
AI
Artificial Intelligence

- **ИИ** – термин скорее из сферы маркетинга
- Часто предпочтение отдаётся другому термину, более корректно отражающему сферу – Наука о данных, Data Science



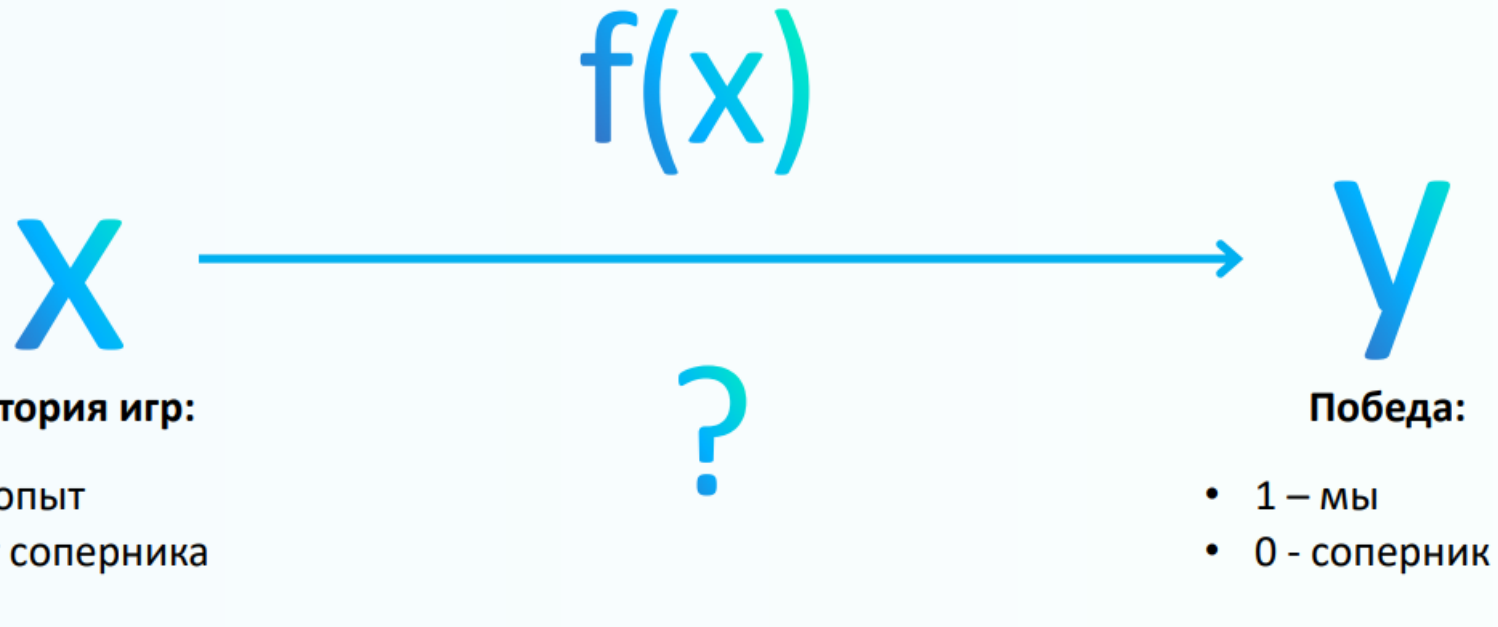
Задачи ИИ

- Пример: отзывы в интернете



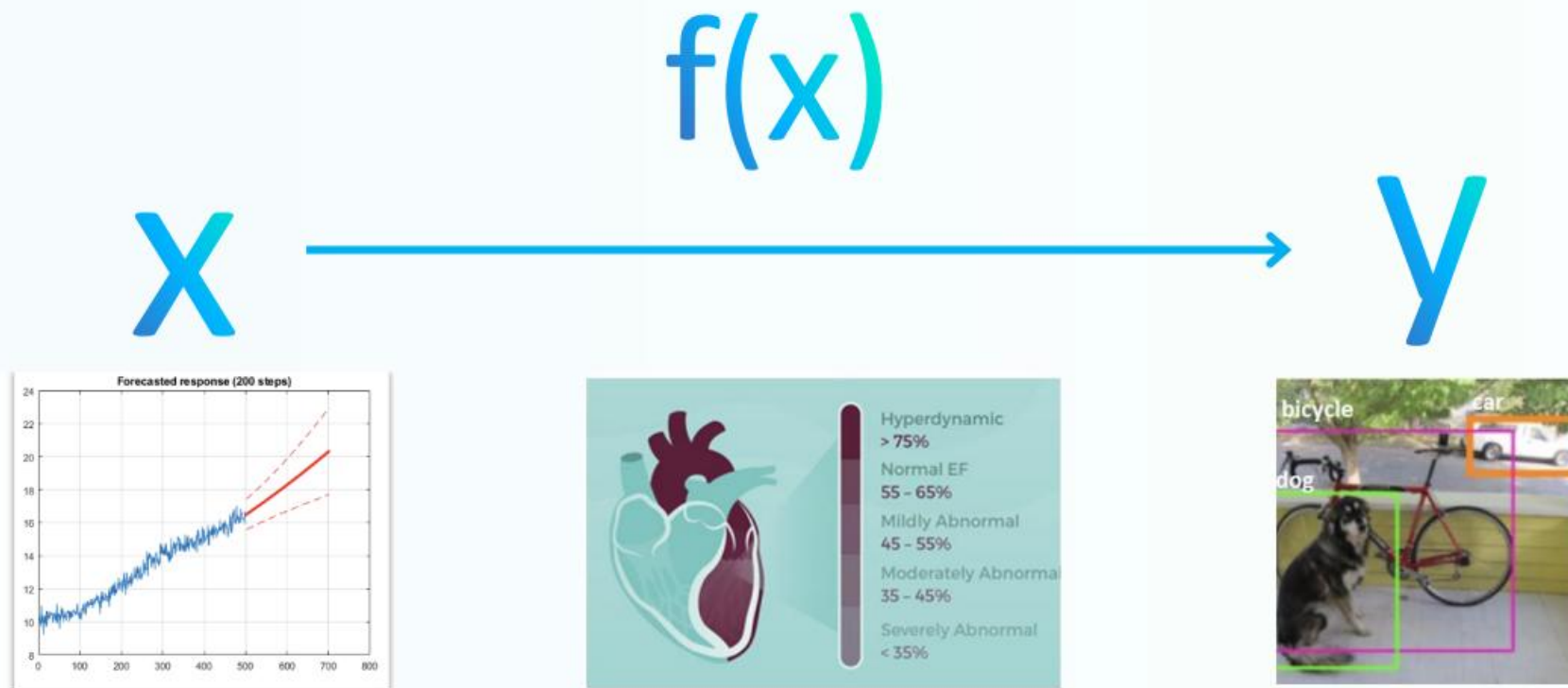
Задачи ИИ

- Пример: победа в игре



Задачи ИИ

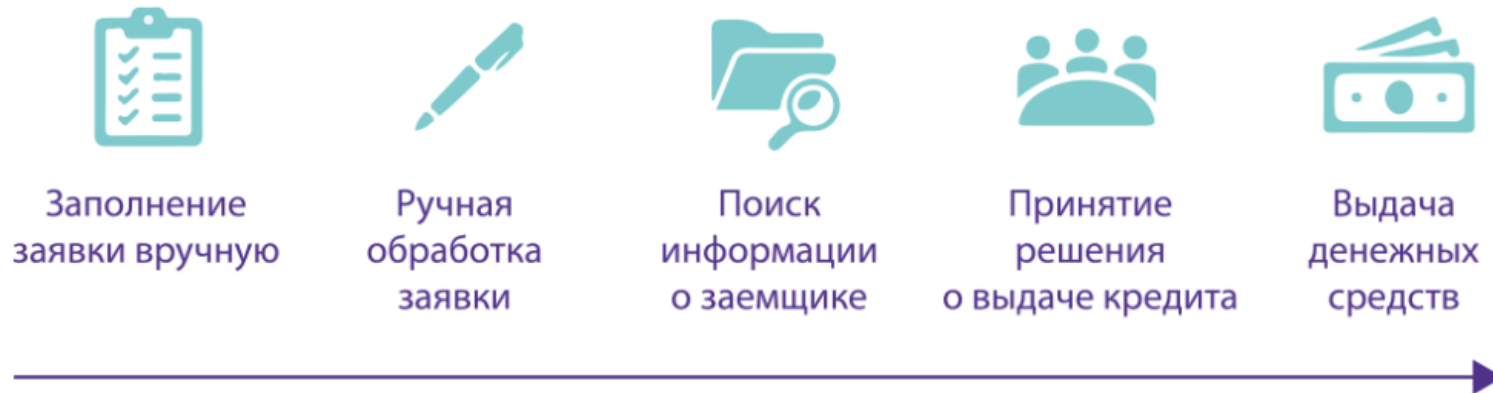
- Больше сложных зависимостей



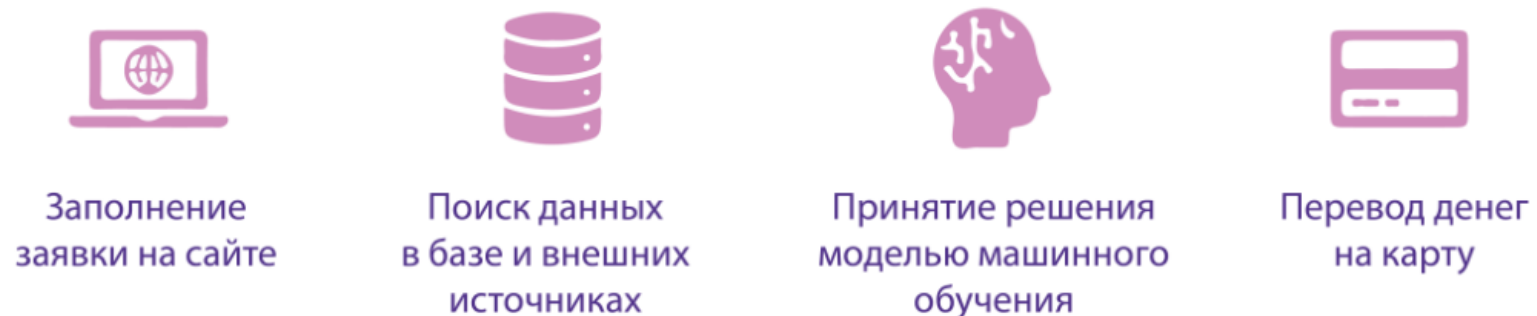
Примеры классических задач

• Кредитный скоринг

Процесс выдачи кредита в 90-х годах



Процесс выдачи кредита в наше время



Примеры классических задач

- Рекомендации фильмов, музыки, товаров



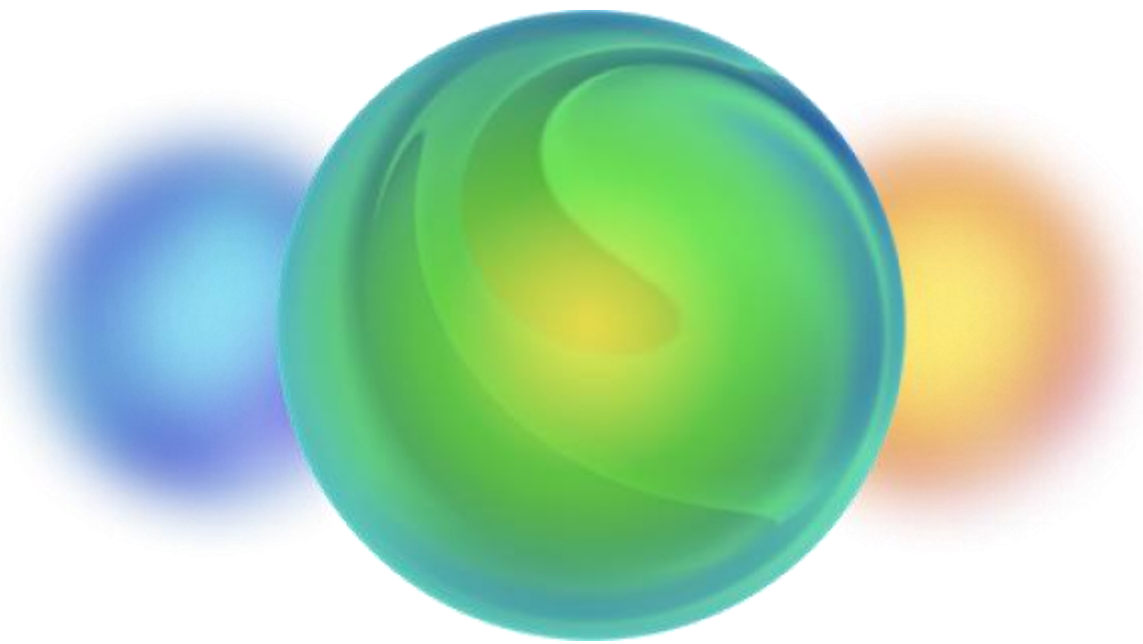
Примеры классических задач

- **Распознавание лиц**



Примеры классических задач

- **Голосовые помощники**



Общие черты задач для ИИ

Очень широкий спектр задач – что между ними общего?

- Все задачи требуют наличия некоторого объема информации или **примеров**, на которых можно **«обучить»** алгоритм.
- **Кредитный скоринг**: исторические данные клиентов.
- **Распознавание лиц**: «размеченные» изображения, на которых вручную выделены и идентифицированы лица. Разметкой занимается человек, а алгоритм пытается подстроиться под поведение человека.

Различия задач для ИИ

Каковы различия?

- Данные могут быть представлены в совершенно разных видах:
 - **Табличные данные** по строкам содержат разные наблюдения (например, клиентов) и по столбцам характеристики этих наблюдений. Один из столбцов — это целевая переменная (например, доход клиента), которую мы хотим предсказать по всем остальным столбцам (другим характеристикам клиента)
 - **Изображения** обычно представлены в виде трехмерных тензоров с числами, отражающими яркость пикселей по каналам RGB
 - **Текст** состоит из набора слов, который преобразуется в матричный вид
 - **Звук** представлен в виде последовательности амплитуд звукового сигнала во времени, которую часто преобразуют в формат спектрограммы

Для каждого типа данных придуманы мириады алгоритмов, каждый со своей спецификой. Их развитие и изобретение новых активно продолжается и сейчас.

Различия задач для ИИ

Каковы различия?

- Данные могут быть представлены в совершенно разных видах:
 - **Табличные данные** по строкам содержат разные наблюдения (например, клиентов) и по столбцам характеристики этих наблюдений. Один из столбцов — это целевая переменная (например, доход клиента), которую мы хотим предсказать по всем остальным столбцам (другим характеристикам клиента)
 - **Изображения** обычно представлены в виде трехмерных тензоров с числами, отражающими яркость пикселей по каналам RGB
 - **Текст** состоит из набора слов, который преобразуется в матричный вид
 - **Звук** представлен в виде последовательности амплитуд звукового сигнала во времени, которую часто преобразуют в формат спектрограммы

Для каждого типа данных придуманы мириады алгоритмов, каждый со своей спецификой. Их развитие и изобретение новых активно продолжается и сейчас.

Понятие ИИ



Без чего невозможно создать модель ИИ?

$$x \xrightarrow[y \approx f(x)]{f(x)} y$$

Понятие ИИ



Без чего невозможно создать модель ИИ?

- Математический аппарат

$$x \xrightarrow[y \approx f(x)]{f(x)} y$$

Понятие ИИ



Без чего невозможно создать модель ИИ?

- Математический аппарат
- Вычислительные мощности

$$x \xrightarrow[y \approx f(x)]{f(x)} y$$

Понятие ИИ



Без чего невозможно создать модель ИИ?

- Математический аппарат
- Вычислительные мощности
- Данные

$$x \xrightarrow[y \approx f(x)]{f(x)} y$$

Современная типизация ИИ

Слабый/узкий ИИ

(Artificial Narrow Intelligence, ANI)

- Одна конкретная задача
- Обучен этой задаче
- Не осознает самого себя
- Не может мыслить независимо

Слабый
(weak)



Сильный
(strong)



Сильный/общий ИИ

(Artificial General Intelligence, AGI)

- Множество задач
- Легко адаптируется, ищет новые знания, планирует будущее
- Самосознание
- Саморефлексия
- Способность мыслить

Современная типизация ИИ

Слабый/узкий ИИ

(Artificial Narrow Intelligence, ANI)

- Одна конкретная задача
- Обучен этой задаче
- Не осознает самого себя
- Не может мыслить независимо

Слабый
(weak)



Сильный
(strong)



Сильный/общий ИИ

(Artificial General Intelligence, AGI)

- Множество задач
- Легко адаптируется, ищет новые знания, планирует будущее
- Самосознание
- Саморефлексия
- Способность мыслить

Супер-ИИ – (Artificial Super Intelligence, ASI) - гипотетический ИИ, который сможет не только воспроизводить максимум способностей человека, но и даже превзойти его. Верящие в ASI считают, что он обретет силу проникновения в мысли и чувства человека с тем, чтобы подчинить его своей воле.



Слабый ИИ

- Слабый ИИ — это ИИ в сегодняшнем понимании.
- ИИ считают слабым, когда машина может справляться только с ограниченным набором отдельных задач лучше человека. Именно на данной стадии сейчас находится тот ИИ, с которым мы с вами сталкиваемся повседневно.
- Слабый ИИ имитирует человеческое познание. Он может принести пользу обществу, автоматизируя трудоемкие задачи и анализируя данные способами, которые людям порой не под силу или чрезвычайно трудозатратны.

Слабый ИИ

Слабому ИИ не хватает человеческого **сознания**, хотя временами он может это имитировать.

Классической иллюстрацией слабого ИИ является мысленный эксперимент Джона Сирла «**Китайская комната**».



Слабый ИИ

- Узкие системы ИИ не обладают общим интеллектом; у них есть свой, "особый", область применения которого сильно ограничена.

Пример: Искусственный интеллект, который знает, как проехать из точки А в точку Б, обычно не способен вызвать Вас на игру в шахматы. Точно так же ИИ, который может притвориться, что говорит с вами по-китайски, вероятно, не сможет подмести ваши полы.

- Такой ИИ может работать в режиме реального времени, но он извлекает информацию лишь из ограниченного набора данных.

Слабый ИИ

Конечно, вряд ли такой ИИ способен на порабощение человечества.

Но все же он уже может превзойти человека — еще в далеком 1997 году машина Deep Blue от компании IBM сумела обыграть мирового чемпиона по шахматам — Гарри Каспарова.



Понятие МО



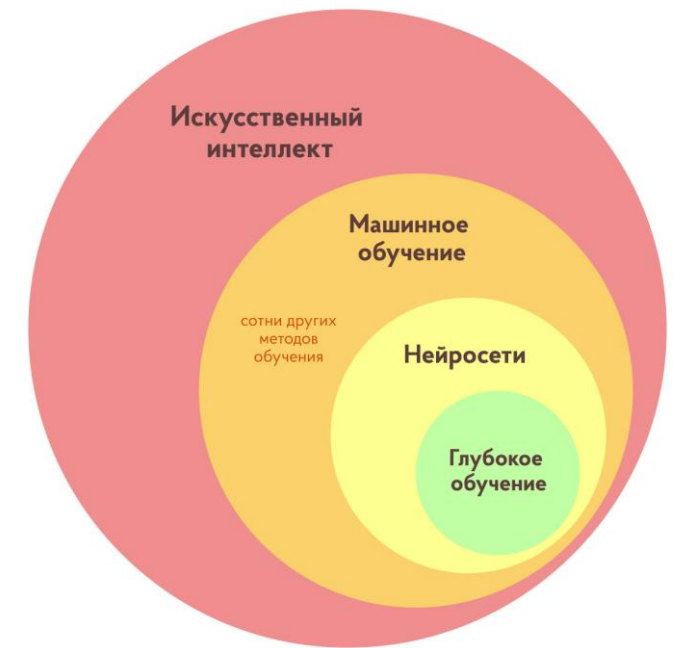
Что понимается под термином
«Машинное обучение»?
(МО, Machine Learning, ML)



Машинное обучение

- **Машинное обучение (МО)** – это подмножество ИИ, фокусирующееся на создании алгоритмов, которые могут обучаться и улучшаться на основе данных.
- Традиционное программирование: все действия прописаны вручную
- Системы машинного обучения: модель МО анализирует данные, выявляет закономерности и строит предсказания или принимает решения на основе новых данных.

МО – это один из основных инструментов, с помощью которого реализуется ИИ.





Машинное обучение

- В МО используются различные методы, включая нейронные сети и другие статистические и математические техники, чтобы научить компьютеры решать конкретные задачи, например, классификацию изображений или прогнозирование.

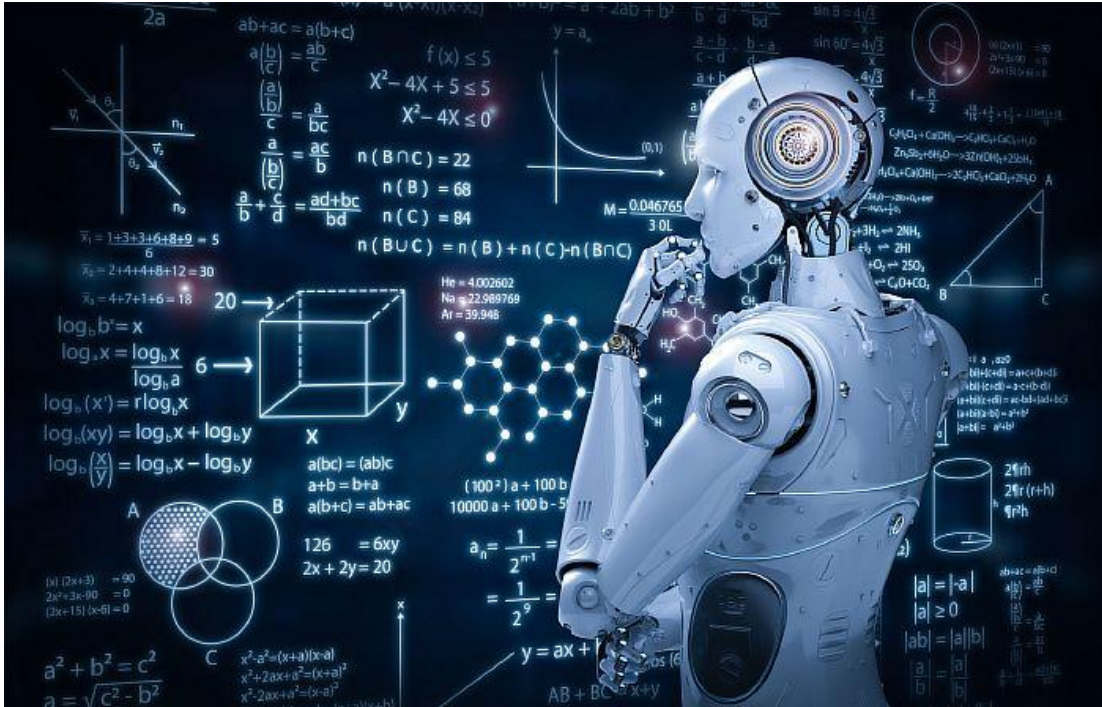
Старый подход

```
function register()
{
    if (!empty($_POST)) {
        $msg = '';
        if ($_POST['user_name']) {
            if ($_POST['user_password_new']) {
                if ($_POST['user_password_new'] === $_POST['user_password_repeat']) {
                    if (strlen($_POST['user_password_new']) > 5) {
                        if (strlen($_POST['user_name']) < 65 && strlen($_POST['user_name']) > 1) {
                            if (preg_match('/^[a-z\d]{2,64}$/i', $_POST['user_name'])) {
                                $user = read_user($_POST['user_name']);
                                if (!isset($user['user_name'])) {
                                    if ($_POST['user_email']) {
                                        if (strlen($_POST['user_email']) < 65) {
                                            if (filter_var($_POST['user_email'], FILTER_VALIDATE_EMAIL)) {
                                                create_user();
                                                $_SESSION['msg'] = 'You are now registered so please login!';
                                                header('Location: ' . $_SERVER['PHP_SELF']);
                                                exit();
                                            } else $msg = 'You must provide a valid email address';
                                        } else $msg = 'Email must be less than 64 characters';
                                    } else $msg = 'Email cannot be empty';
                                } else $msg = 'Username already exists';
                            } else $msg = 'Username must be only a-z, A-Z, 0-9';
                        } else $msg = 'Username must be between 2 and 64 characters';
                    } else $msg = 'Password must be at least 6 characters';
                } else $msg = 'Passwords do not match';
            } else $msg = 'Empty Password';
        } else $msg = 'Empty Username';
        $_SESSION['msg'] = $msg;
    }
    return register_form();
}
```



«Используем миллион правил, которые учитывают все особенности задачи»

Новый подход



The image shows a futuristic robot with a glowing head, surrounded by various mathematical and scientific diagrams. The diagrams include:

- A coordinate system with a line and points $\alpha_1, \alpha_2, \alpha_3$.
- Algebraic identities: $ab+ac = a(b+c)$, $\frac{a}{\frac{b}{c}} = \frac{ab}{c}$, $\frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a}{bc}$, $\frac{a}{\frac{b}{c}} = \frac{ac}{b}$, $\frac{\frac{a}{b}}{\frac{c}{d}} = \frac{ad+bc}{bd}$.
- Quadratic equations: $f(x) \leq 5$, $X^2 - 4X + 5 \leq 5$, $X^2 - 4X \leq 0$.
- Probability: $n(B \cap C) = 22$, $n(B) = 68$, $n(C) = 84$, $n(B \cup C) = n(B) + n(C) - n(B \cap C)$.
- Logarithms: $\log_b b^x = x$, $\log_b x = \log_b x$, $\log_b a$, $\log_b(x^r) = r \log_b x$, $\log_b(xy) = \log_b x + \log_b y$, $\log_b\left(\frac{x}{y}\right) = \log_b x - \log_b y$.
- Geometry: A 3D cube with dimensions 20, 6, and X. A Venn diagram with sets A, B, and C. A cylinder with height 2r and radius r.
- Chemical formulas: $CH_4 + Cl_2 \rightarrow CH_3Cl + HCl$, $ZnSb_2 + H_2O \rightarrow Zn(OH)_2 + Sb_2H_3$, $4O_2 + C_2H_6 \rightarrow 2C_2H_4O + 2H_2O$.
- Physics: $M = \frac{0.046765}{3.0L}$, $He = 4.002602$, $Na = 22.989769$, $Ar = 39.948$.
- Algebra: $a^2 + b^2 = c^2$, $a = \sqrt{c^2 - b^2}$, $x^2 - a^2 = (x+a)(x-a)$, $x^2 + 2bx + a^2 = (x+a)^2$, $x^2 - 2bx + a^2 = (x-a)^2$.
- Linear equations: $126 = 6xy$, $2x + 2y = 20$, $a_n = \frac{1}{2^{n-1}}$, $= \frac{1}{2^{m-1}}$, $= \frac{1}{2^9}$.
- Other: $|a| = |-a|$, $|a| \geq 0$, $|ab| = |a||b|$, $\frac{a}{\frac{b}{c}} = \frac{ac}{b}$, $AB + BC = AC$.

«Зальём обучающийся алгоритм данными и он сам найдёт в них закономерности»

Нейронные сети

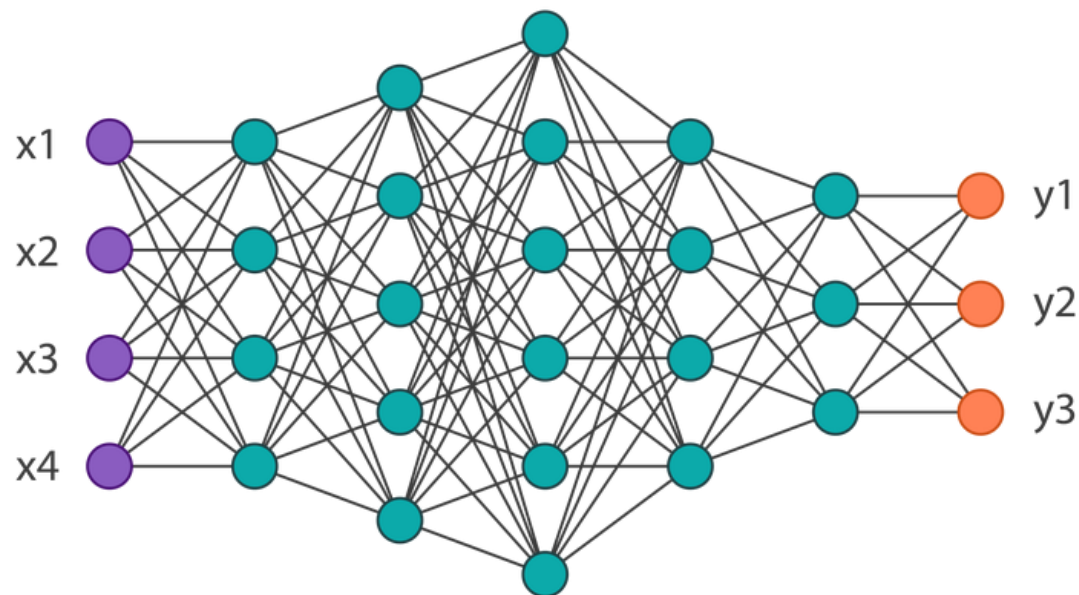
Нейронные сети

Artificial Neural networks, ANN

Нейронные сети - это класс моделей машинного обучения, вдохновленных структурой и функцией нейронов в человеческом мозге.

Они состоят из искусственных нейронов, объединенных в слои, и используются для обработки данных и извлечения признаков.

Нейронные сети могут быть использованы в различных задачах, от распознавания образов до обработки текста и много чего еще.



Глубокое обучение

Глубокое обучение Deep Learning, DL

Глубокое обучение - это подраздел машинного обучения, который сосредотачивается на использовании глубоких нейронных сетей с множеством слоев для решения сложных задач.

Глубокое обучение обычно требует большого количества данных и вычислительных ресурсов, но может достичь выдающихся результатов в различных областях применения.



ChatGPT – модель глубокого обучения

Виды машинного обучения



- По степени привлечения учителя:
 - Обучение с учителем (supervised learning)
 - Классификация (classification)
 - Регрессия (regression)
 - Обучение без учителя (unsupervised learning)
 - Кластеризация (clustering)
 - Обучение с частичным привлечением учителя (semi-supervised learning)

Основные термины и понятия

Данные (Data)

Данные - информация, используемая для обучения модели.

Она может включать в себя:

- **Признаки** (характеристики, фичи, features, обозначение - X)
- **Целевую переменную** (целевое значение, target, обозначение - Y) (в задачах обучения с учителем).

Площадь квартиры (X1, feature)	Количество комнат (X2, feature)
73	2
150	4
34	1
101	3

Данные, содержащие только признаки

Площадь квартиры (X1, feature)	Количество комнат (X2, feature)	Цена квартиры (Y, label, target)
73	2	5.5
150	4	15.8
34	1	4.2
101	3	10.9

Данные, содержащие признаки и целевую переменную

Основные термины и понятия



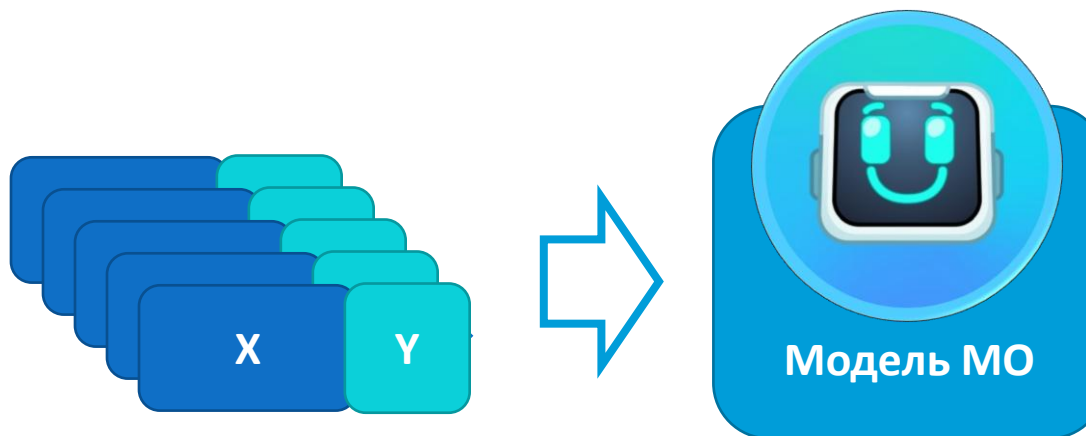
Обучение с учителем

“Вот примеры, как нужно решить задачу. Научись на них ее решать”

Модель обучается на основе **размеченных данных** (наборе данных, содержащих явно заданное значение целевой переменной).

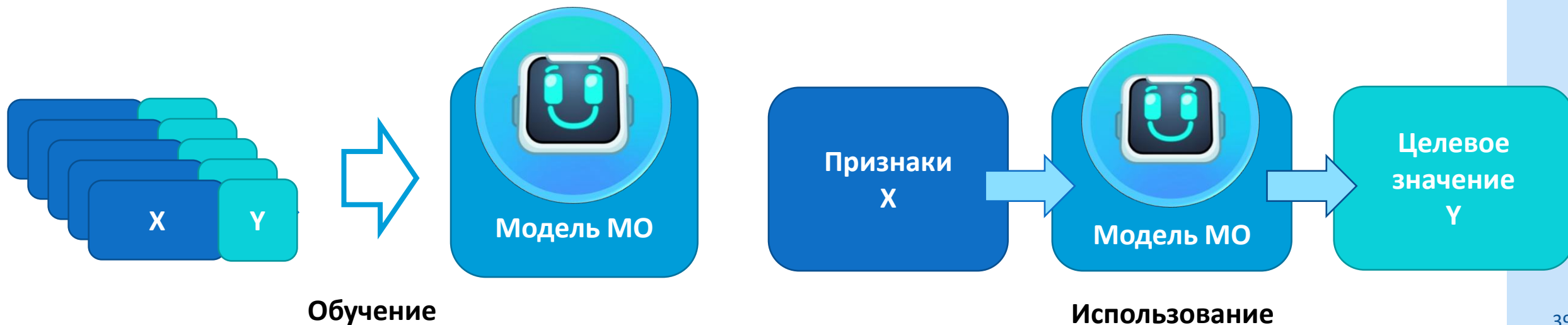
В этом типе задачи каждый обучающий пример представляет собой пару: **входные данные** (признаки) и соответствующее этим данным **значение выходной метки** (целевая переменная).

Задача модели состоит в том, чтобы научиться предсказывать выходные метки на основе входных данных.



Обучение с учителем, основные моменты

- **Размеченные данные:** Для обучения модели требуются данные, для которых известны правильные ответы или метки.
- **Целевая переменная:** Обучающие примеры включают в себя целевую переменную, которую модель должна предсказать.
- **Обучение на основе примеров:** Модель обучается на основе обучающих примеров, применяя различные алгоритмы и методы для поиска связей и закономерностей между входными данными и целевой переменной.
- **Цель - обобщение:** Основная цель обучения с учителем - научить модель обобщать знания с обучающих данных на новые, ранее не виденные данные. Модель должна способно предсказывать метки для новых примеров.

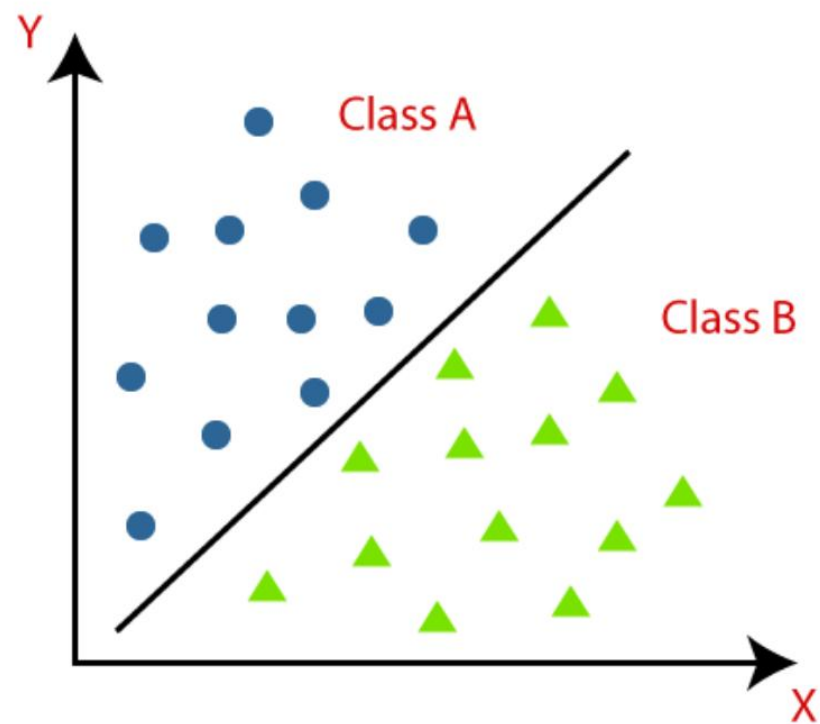


Задачи обучения с учителем: Классификация

Классификация – предсказание категории или класса для входных данных.

Целевое значение Y принимает значение одного из заранее определённых классов.

- **Бинарная классификация:** В задачах бинарной классификации метка может принимать одно из двух значений, например, 0 или 1, Да или Нет, Положительный или Отрицательный и т.д.
- **Многоклассовая классификация:** В многоклассовой классификации метка может относиться к одному из нескольких классов или категорий. Например, при классификации изображений метка может указывать на тип объекта (кошка, собака, автомобиль и так далее)

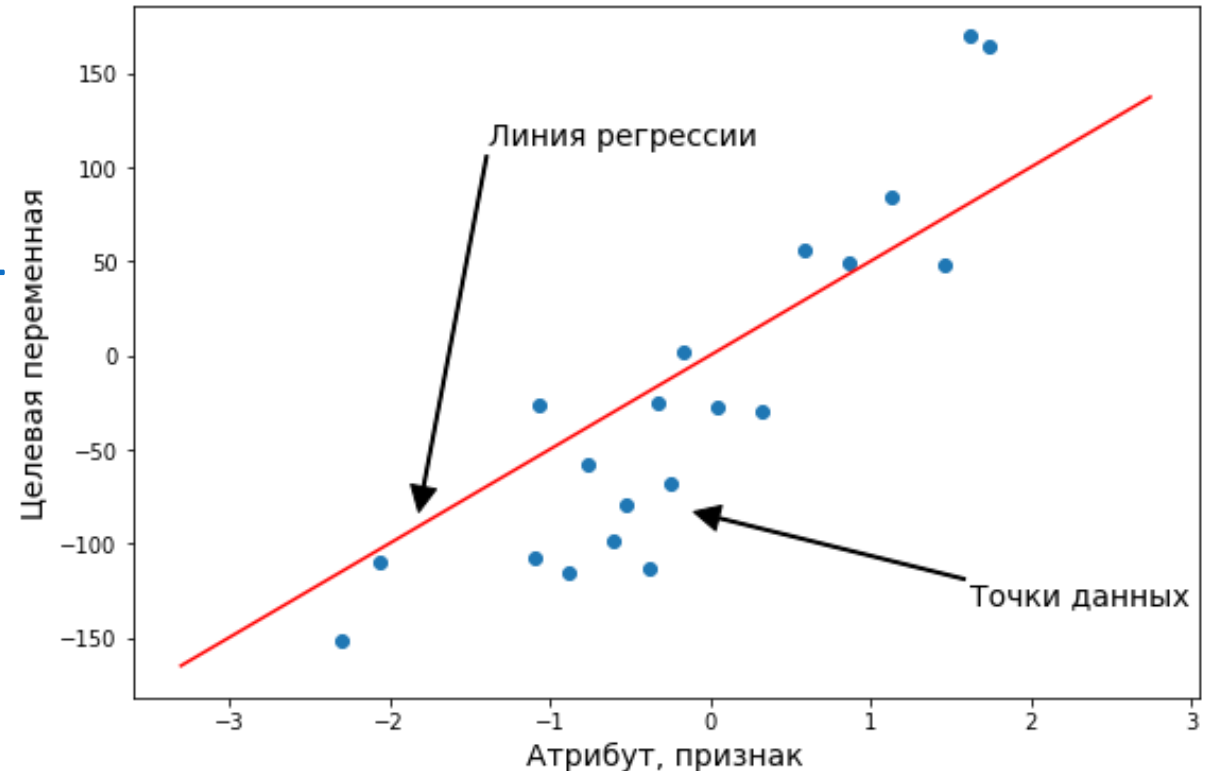


Задачи обучения с учителем: Регрессия

Регрессия – предсказание числового значения по входным данным.

Целевое значение Y представляет собой числовое значение, которое модель пытается предсказать.

Например, при прогнозировании цены дома целевое значение может быть числом, представляющим стоимость.



Выходная величина расположена на отрезке значений

Целевое значение

- **Классификация**

Целевое значение категориальное – метка дискретных категорий или классов.

Например, в задаче классификации изображений целевая переменная может указывать на категорию объекта на изображении (например, "кошка", "собака", "автомобиль").

- **Регрессия**

Целевое значение численное – конкретное числовое значение.

Модель стремится предсказать точное численное значение.

Примеры численных целевых переменных: цена товара, температура, доход, количество продаж и так далее.

Обучение без учителя

“Вот данные, в которых нужно найти зависимости. Попробуй их найти”

Модель обучается на основе **неразмеченных данных** (в наборе данных **отсутствует значение целевой переменной**).

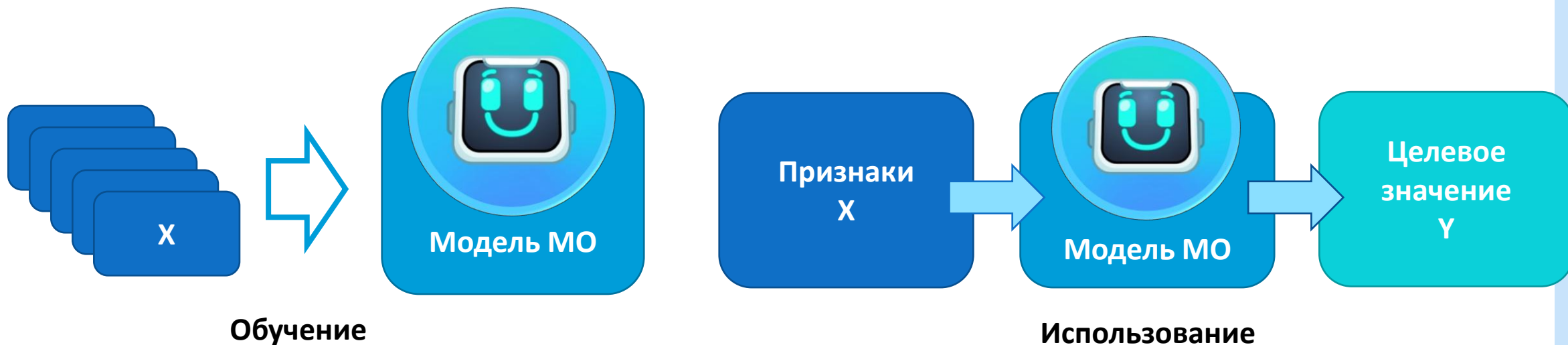
Модель пытается извлечь скрытую структуру или закономерности из входных данных **без предварительной информации о правильных ответах**.

Этот тип задач машинного обучения позволяет модели самостоятельно выявлять закономерности, группировать данные и создавать представления о данных.



Обучение с учителем, основные моменты

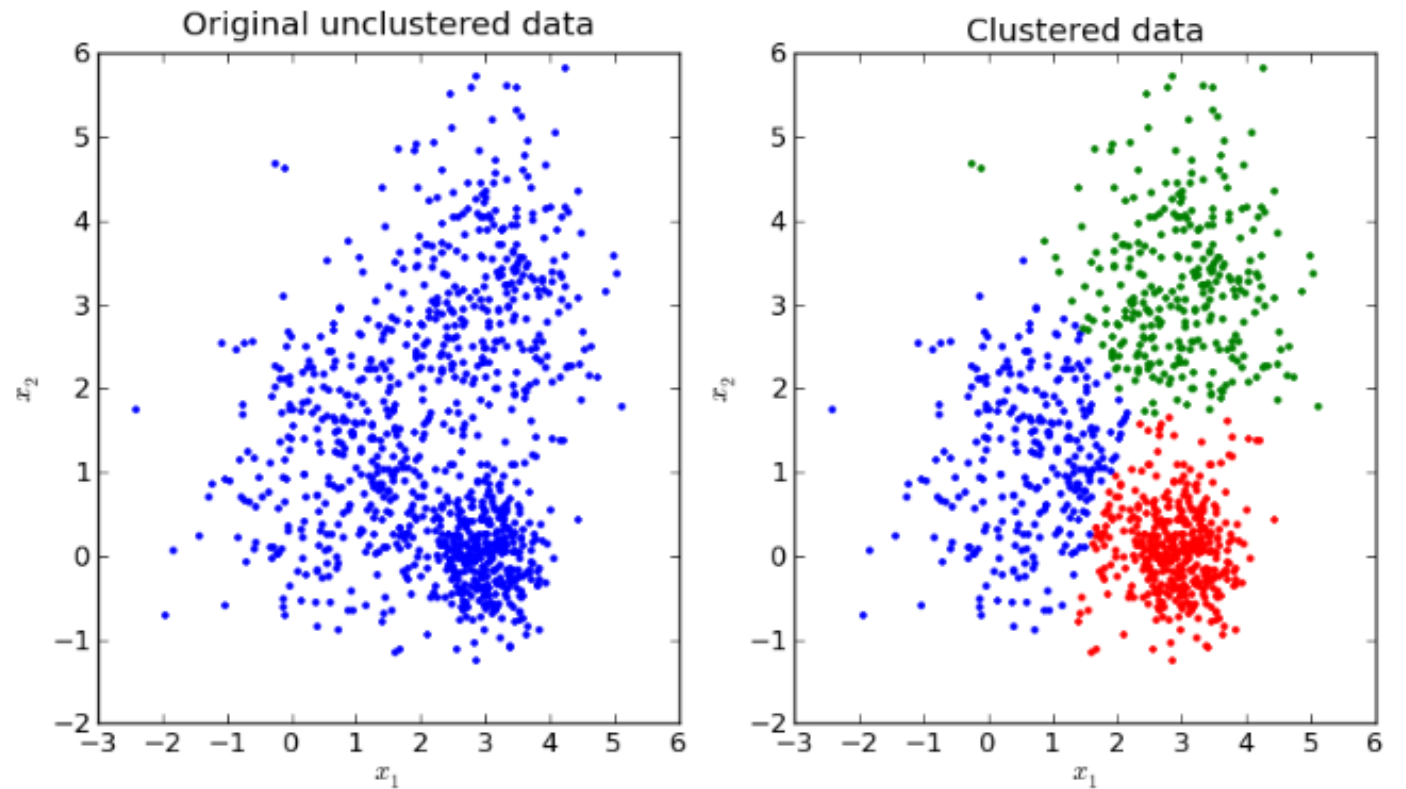
- **Отсутствие целевой переменной** в обучающих данных. В задачах обучения с учителем у нас есть целевая переменная, которую модель пытается предсказать. В обучении без учителя целевая переменная отсутствует.
- **Группировка и структура данных:** Обучение без учителя ориентировано на поиск структуры или закономерностей в данных. Модель стремится выделить группы схожих объектов или другие характеристики данных, такие как кластеры или сниженная размерность.
- **Цель - извлечение информации:** Главной целью обучения без учителя является извлечение полезной информации из данных. Модель создает представление данных, которое может помочь в анализе, визуализации, сегментации или других задачах.
- **Обучение на основе характеристик :** В обучении без учителя модель основывается на признаках данных, чтобы выявить их структуру или закономерности.



Задачи обучения без учителя: Кластеризация

Кластеризация – это группировка объектов в **кластеры** (группы схожих объектов) на основе их сходства.

Кластеризация позволяет выявить внутренние структуры в данных и определить, какие объекты близки по своим характеристикам.



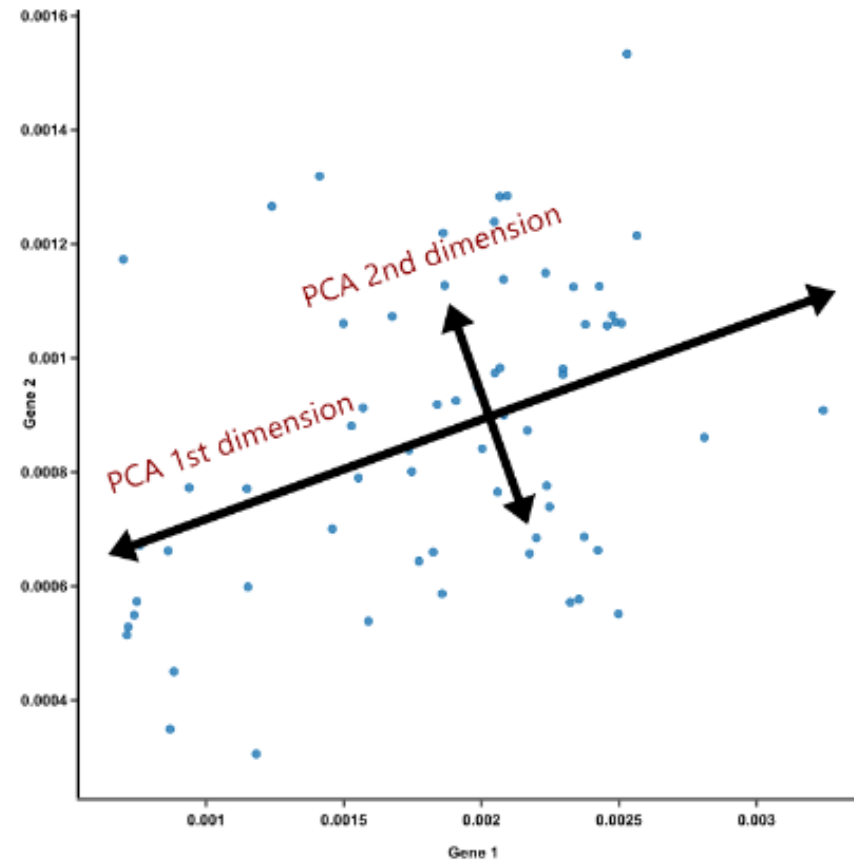
Отличие кластеризации и классификации

Классификация	Кластеризация
<p>Цель: <u>присвоение объектов или данных одной из заранее определенных категорий или классов.</u></p> <p>В классификации у нас есть явные метки или метки классов, которые мы пытаемся предсказать для новых данных на основе характеристик объектов.</p> <p>Целевая переменная: В задачах классификации <u>существует явная целевая переменная</u>, которую модель пытается предсказать для новых данных. Целевая переменная состоит из классов или категорий.</p>	<p>Цель: <u>группировка схожих объектов или точек данных на основе их характеристик без заранее известных категорий или меток.</u></p> <p>В кластеризации мы пытаемся найти структуру в данных, выделяя группы или кластеры объектов, которые имеют схожие характеристики или свойства.</p> <p>Целевая переменная: В задачах кластеризации <u>нет заданной заранее целевой переменной</u> или меток классов. <u>Модель самостоятельно определяет структуру кластеров в данных.</u></p>

Задачи обучения без учителя: Снижение размерности

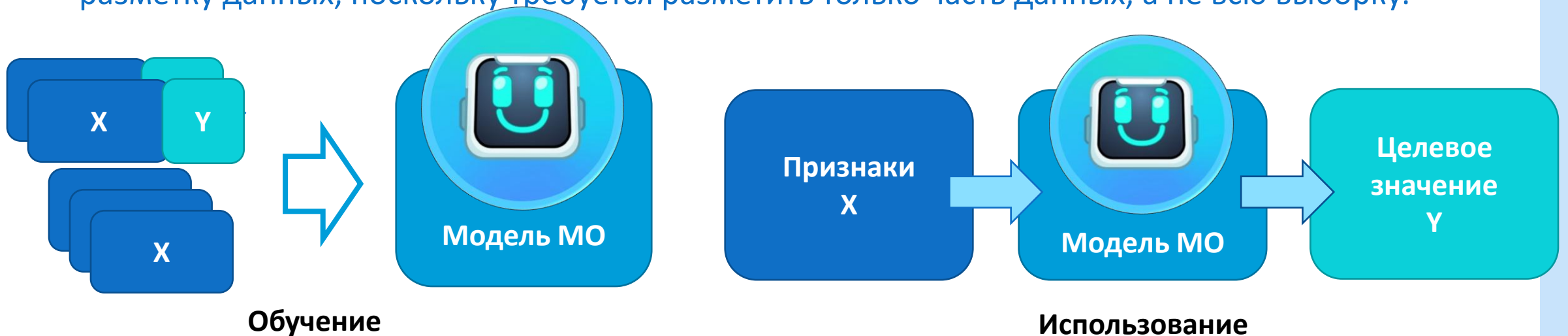
Снижение размерности – это уменьшение размерности данных путем проекции на более низкоразмерное пространство. При этом стремятся сохранить максимальное количество информации.

Примером является метод главных компонент (Principal Component Analysis, PCA).



Обучение с частичным привлечением учителя

- **Использование размеченных данных:** Модель использует размеченные данные для обучения, чтобы научиться предсказывать целевую переменную.
- **Использование неразмеченных данных:** Дополнительно к размеченным данным, модель также использует немаркированные данные для извлечения структуры данных или для улучшения обобщения. Это помогает модели учить более общие закономерности и снижать риск переобучения.
- **Воссоздание разметки неразмеченных данных:** Существует ряд алгоритмов и методов, разработанных специально для обучения с частичным привлечением учителя. Некоторые из них используют методы активного обучения, которые позволяют модели запрашивать разметку для наиболее неопределенных или важных примеров.
- **Снижение нагрузки на разметку данных:** Этот метод позволяет снизить трудозатратность на разметку данных, поскольку требуется разметить только часть данных, а не всю выборку.

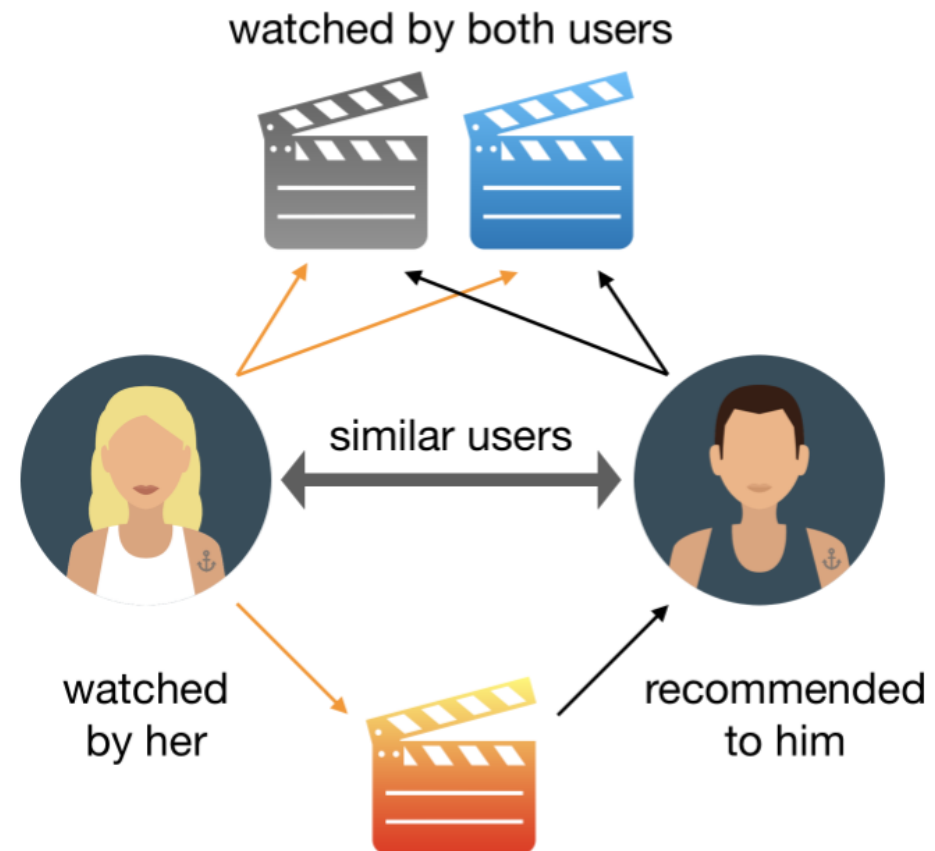


Другие виды машинного обучения:

Рекомендательные системы

Класс программных и алгоритмических инструментов, которые анализируют данные о предпочтениях или поведении пользователей, чтобы предложить им релевантные и персонализированные рекомендации.

Они используются для предсказания, какие товары, услуги, контент или информация могут быть интересны конкретному пользователю.

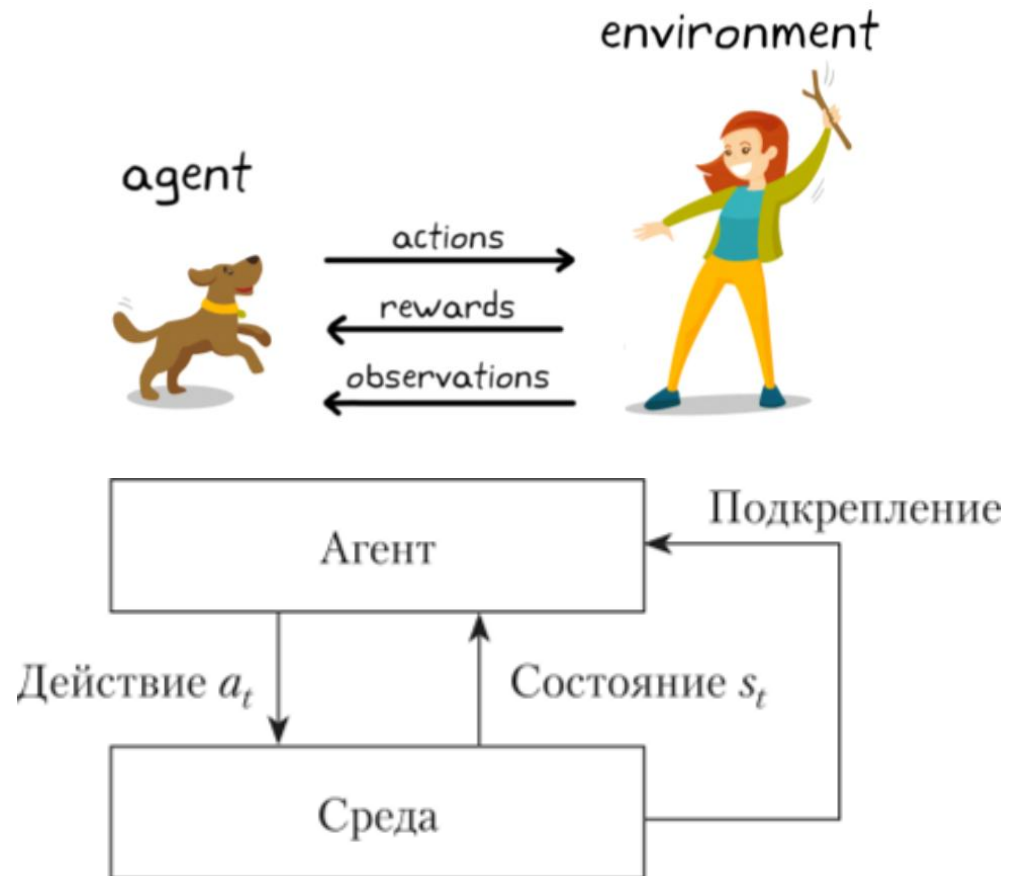


Другие виды машинного обучения:

Обучение с подкреплением

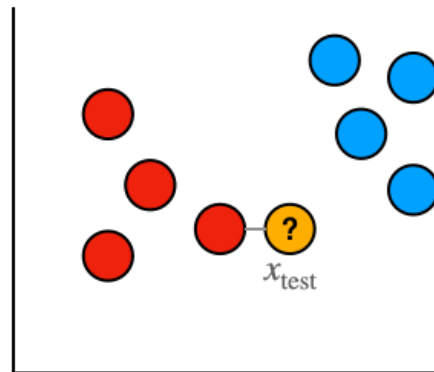
Подход в машинном обучении, в котором агент (или программа) учится принимать последовательность решений, чтобы максимизировать некоторую численную награду (или минимизировать некоторый штраф) в определенной среде.

Этот подход аналогичен обучению, которое происходит у человека или животного, когда они взаимодействуют с окружающей средой, принимают решения и адаптируются на основе получаемого опыта.



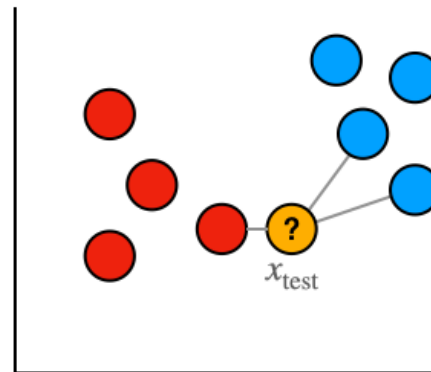
Алгоритм kNN

К-ближайших соседей (K-Nearest Neighbors или просто kNN) — алгоритм классификации и регрессии, основанный на гипотезе компактности, которая предполагает, что расположенные близко друг к другу объекты в пространстве признаков имеют схожие значения целевой переменной



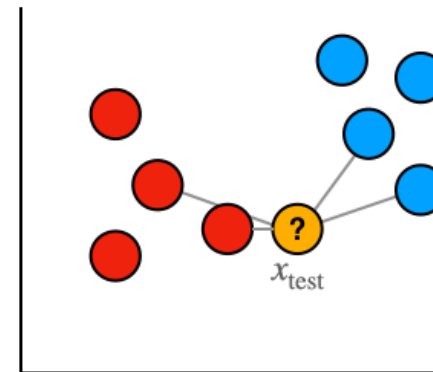
$k = 1$

Nearest point is **red**, so x_{test} classified as **red**



$k = 3$

Nearest points are {**red**, **blue**, **blue**} so x_{test} classified as **blue**



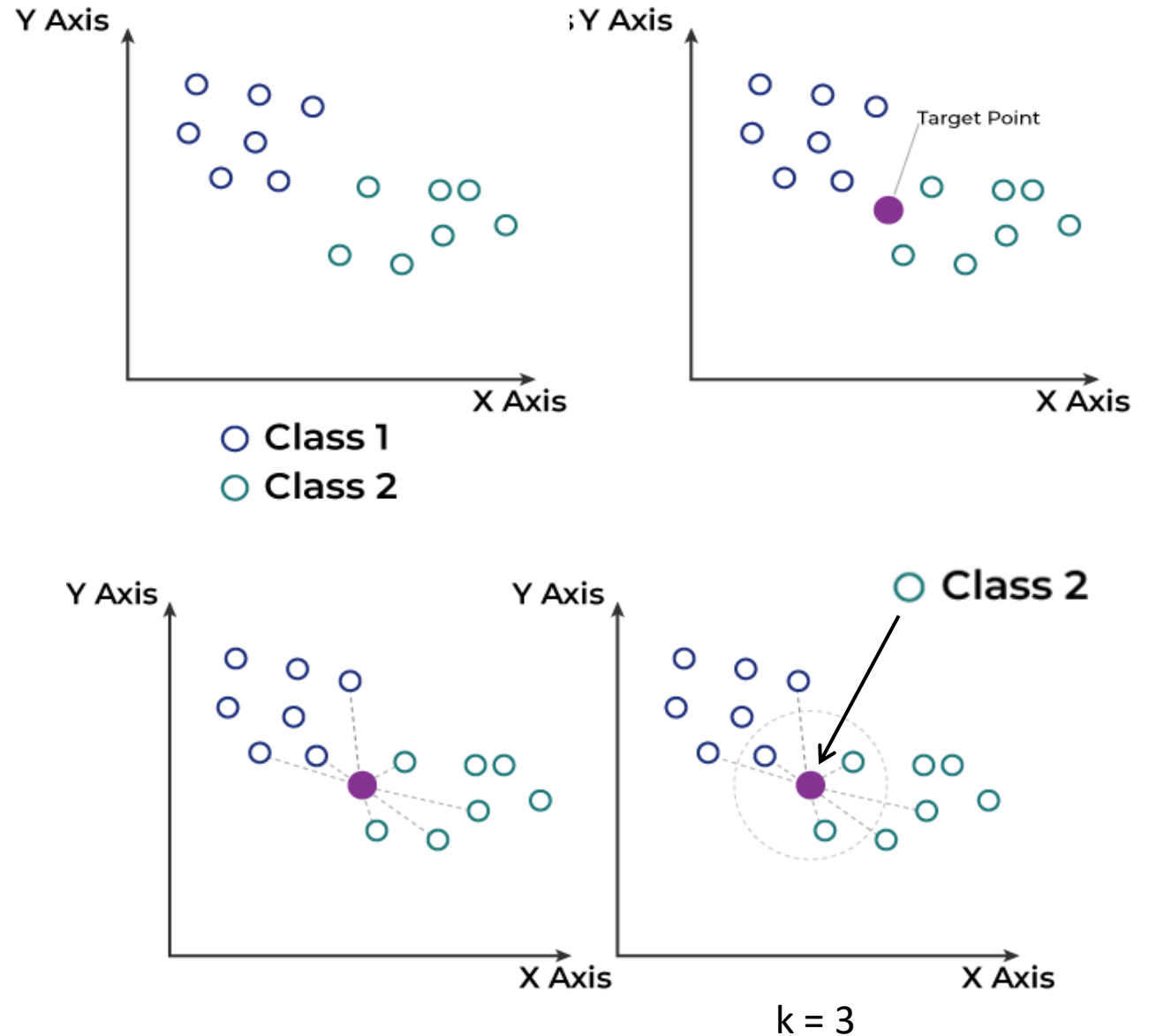
$k = 4$

Nearest points are {**red**, **red**, **blue**, **blue**} so classification of x_{test} is not properly defined

Алгоритм kNN

Принцип работы KNN:

- 1) сначала вычисляется расстояние между тестовым и всеми обучающими образцами;
- 2) далее из них выбирается k-ближайших образцов (соседей), где число k задаётся заранее;
- 3) итоговым прогнозом среди выбранных k-ближайших образцов будет мода в случае классификации и среднее арифметическое в случае регрессии;
- 4) предыдущие шаги повторяются для всех тестовых образцов.



Преимущества и недостатки KNN

Преимущества:

- простота в реализации и интерпретации;
- применяется во многих задачах, особенно в рекомендательных системах;
- высокая точность прогнозов при правильном подборе k и метрики расстояния.

Недостатки:

- большое потребление памяти и низкая скорость работы из-за хранения и вычисления расстояний между всеми обучающими и тестовыми образцами;
- чувствительность к выбросам и шуму, а также к несбалансированным классам в данных;
- при большом количестве признаков может возникнуть проблема совпадения метрической и смысловой близости объектов.

Искусственный Интеллект

Спасибо за внимание!