



ТИНЬКОФФ

Лекция 5

Машинное обучение

Начало



Постановка задачи в общем виде

X

множество объектов
(или их описаний)

Y

множество значений
целевой переменной

дано:

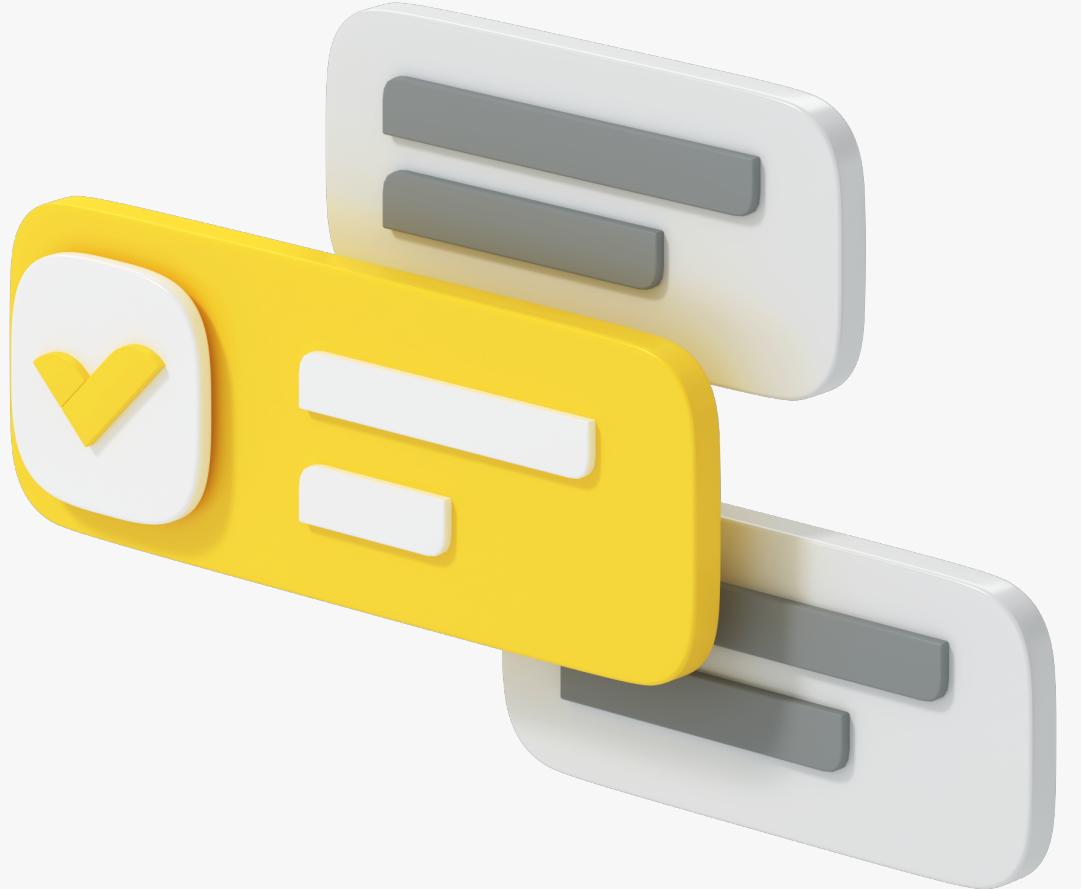
$\{x_1, \dots, x_l\} \subset X$ - обучающая выборка

$y_i = y(x_i), i = 1, \dots, l$ - известные ответы

Найти:

$a: X \rightarrow Y$ — алгоритм, строящий
приближение $y(x)$ на всем
множестве X

Классификация задач по типу обучения



01

Обучение с учителем

Существует некоторая выборка, на которой известны целевые переменные

02

Обучение без учителя

Отсутствуют целевые переменные, требуется что-то сделать с данными

03

Частичное обучение

(semi-supervised learning)

Знаем целевые переменные на некотором небольшом множестве ответов

Классификация задач по типу таргета



01

Классификация

Целевая переменная
дискретна и конечна

02

Регрессия

Целевая переменная
непрерывной величиной

03

**Задачи
ранжирования**

Результатом алгоритма
является некоторое
упорядоченное множество

04

**Обучение
без учителя**

Целевая переменная
отсутствует

Качество модели

функция потерь —

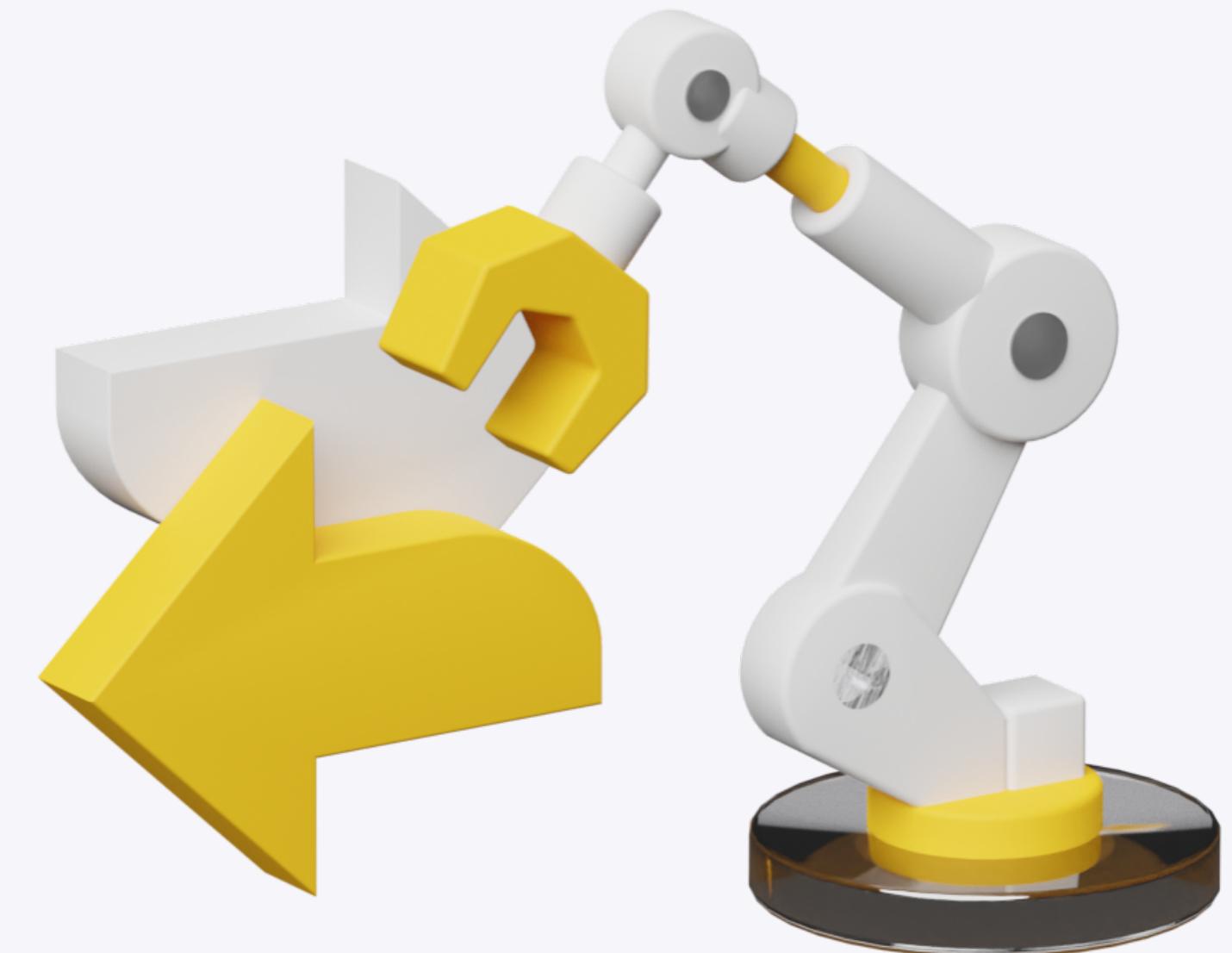
величина ошибки модели a на объекте x из X

Эмпирический риск —

функционал качества модели a на всей
выборке X^l



Качество моделей оценивается
по эмпирическому риску



Примеры функционалов потерь



Средне квадратичная ошибка (MSE)

$$\frac{1}{n} \sum_{i=1}^n (y_i - a(x_i))^2$$



Средне абсолютная ошибка (MAE)

$$\frac{1}{n} \sum_{i=1}^n |y_i - a(x_i)|$$



Средняя ошибка классификации

$$\frac{1}{n} \sum_{i=1}^n [a(x_i) \neq y_i]$$



Кросс-энтропия

$$\sum_{i=1}^n (-y_i \log a(x_i) - (1 - y_i) \log(1 - a(x_i)))$$

Примеры задач



Задача кредитного scoringа



Задача оттока клиентов



Детекция аномалий



Детекция аномалий

Линейная регрессия

Модель

Веса, являющееся статистикой, полученной через минимизацию MSE, являются статистикой максимального правдоподобия.

Поэтому для поиска весов будем минимизировать MSE градиентными методами

Градиент MSE: $-\frac{2}{n}X^T(Y - Xw)$

**В виде функции
многих переменных:**

$$a(w, x) = w_0 + x_1 w_1 + \dots + x_l w_l$$

В матричном виде:

$$a(X) = Xw$$



ТИНЬКОФФ

Он такой один