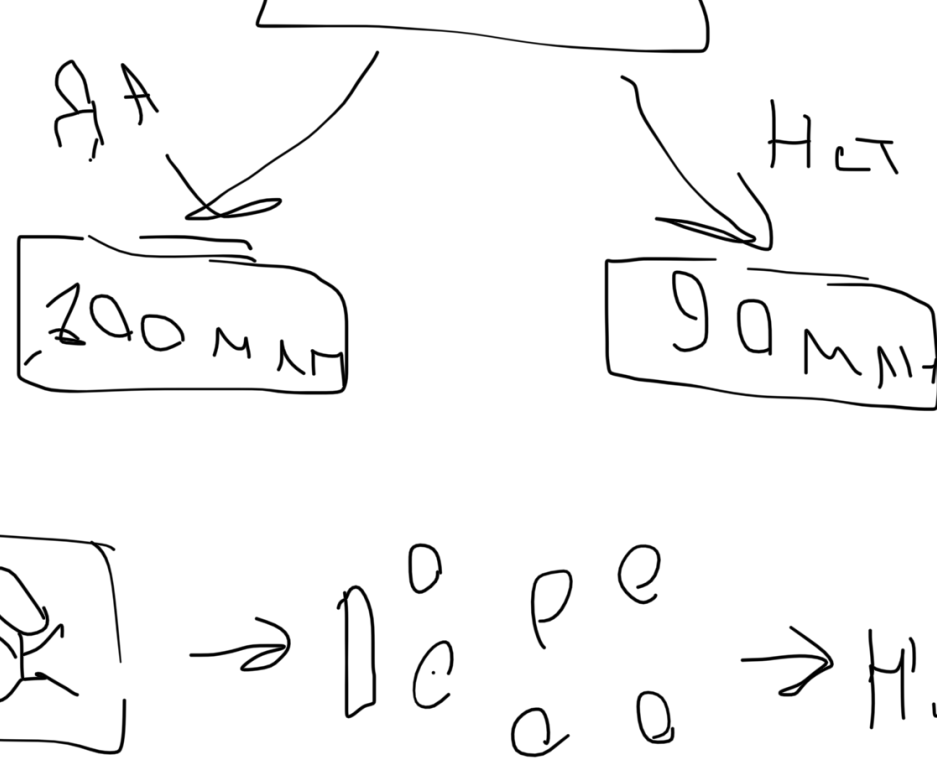


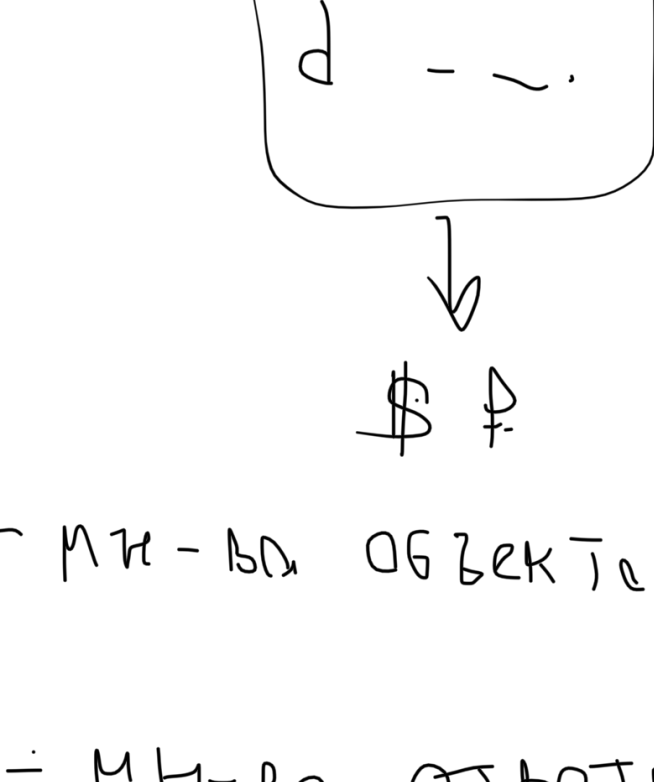
Марк

Елена Кантанистова

$$y = kx + b$$



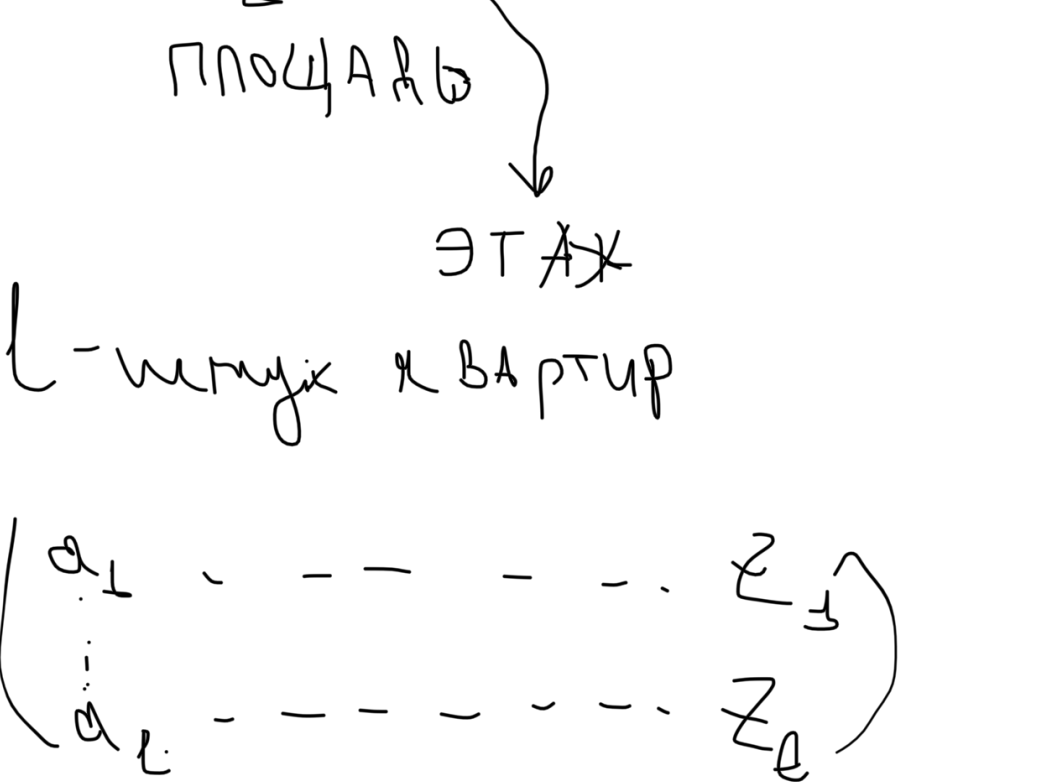
$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \rightarrow \text{Матрица}$$



X - мн-во объектов

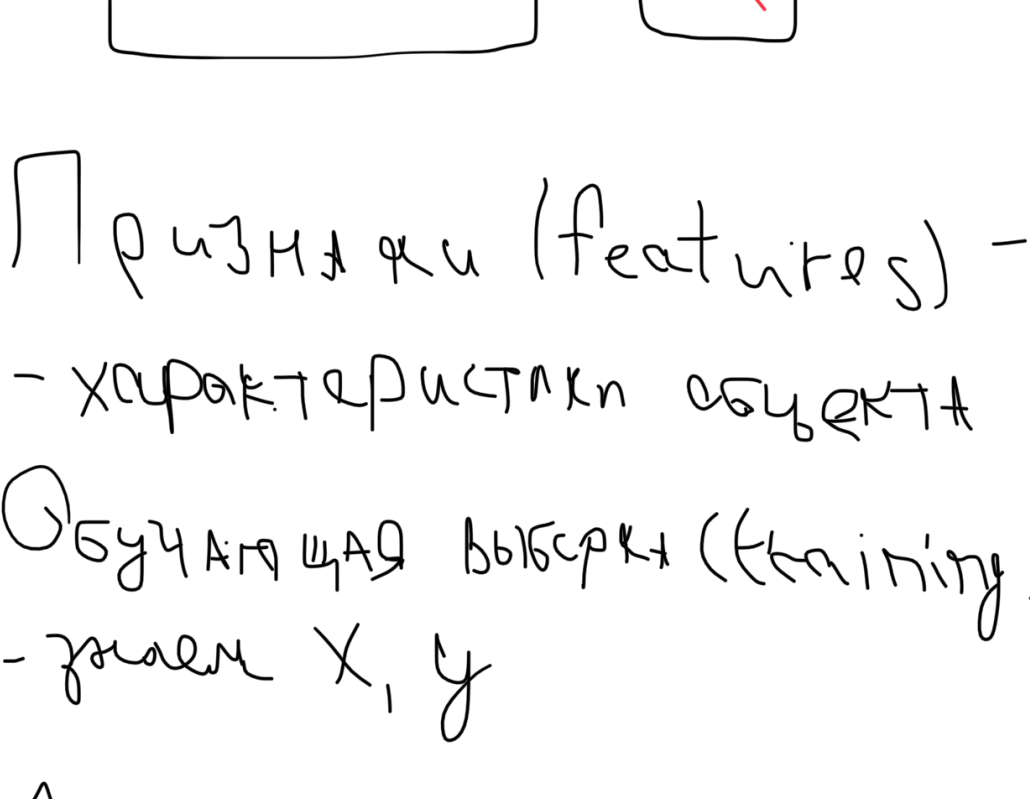
Y - мн-во ответов

$$A: X \rightarrow Y$$



l-мн-во квартир

$$\begin{pmatrix} a_1 & \dots & z_1 \\ \vdots & \ddots & \vdots \\ a_l & \dots & z_l \end{pmatrix}$$



Признаки (features) -

- характеристики объекта

Обучающая выборка (training set) -

- знаем X, y

Виды признаков:

- 1) Числовые (площадь)
- 2) Бинарные (0/1)
- 3) Категориальный (название города)
- 4) Сложные признаки (фото)

Виды данных:

- 1) Текст
- 2) Изображения
- 3) Таблицы (XLS, CSV)
- 4) Звук
- 5) Логг
- 6) Streamer

Виды целевой переменной (y)

1) Классификация:

- a) $y = \{0, 1\}$
- б) $y = \{1, \dots, M\}$
- в) $y = \{0, 1\}^n$ - классификация с пересечениями

2) Регрессия

$$y = \mathbb{R}$$

$$y = \mathbb{R}^n$$

3) Ранжирование:

y - упорядоченное мн-во

4) Кластеризация

y просто нет

5) Ранжирование размерности

6) Оценивание пл-ти

7) Визуализация

Типы обучения:

а) с учителем (y-есть)

б) без (y-нет)

$$a, b, c, d \leftarrow \text{house icon}$$

$$A: \hat{y} = w_0 + w_1 a + w_2 b + \dots$$

$$w_0, \dots, w_n$$

$$A = \{ w_0, w_1, \dots, w_n \}$$

$$w_0, w_1, w_2, w_3, w_4$$

$$Q(A, X) = \frac{1}{L} \sum_i (w_0 + w_1 a + w_2 b + w_3 c + w_4 d - y_i)^2 \rightarrow 0$$

$$\min$$

$$w_0, w_1, w_2, w_3, w_4$$

Обучение с поиском оптимальных

$$w_0, \dots, w_4$$

Метрики качества

Регрессия: MSE, MAE

Классификация:

$$\text{accuracy}(A, X) =$$

$$\frac{1}{L} \sum_i [k(x_i) = y_i]$$

$$\text{if } A(x_i) = y_i : 1$$

$$\text{else} : 0$$

Алгоритм Решения:

- 1) Постановка задачи
- 2) Выделяете признаки
- 3) Формирование выборки
- 4) Выбор функции потерь и метрики качества
- 5) Преобразование данных
- 6) Построение модели
- 7) Оценивание качества