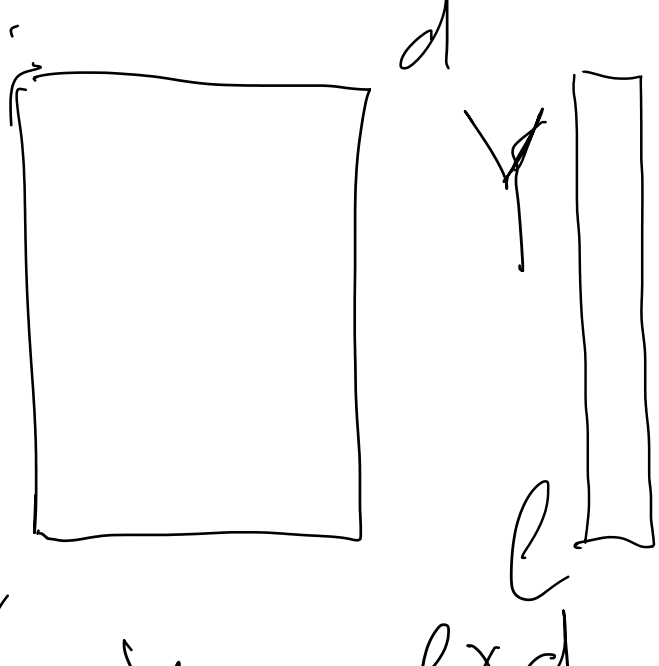
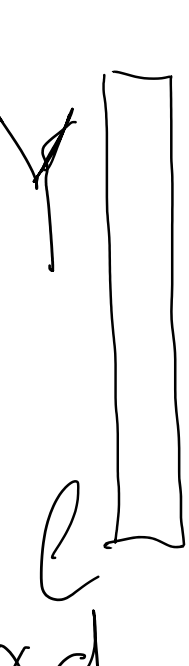


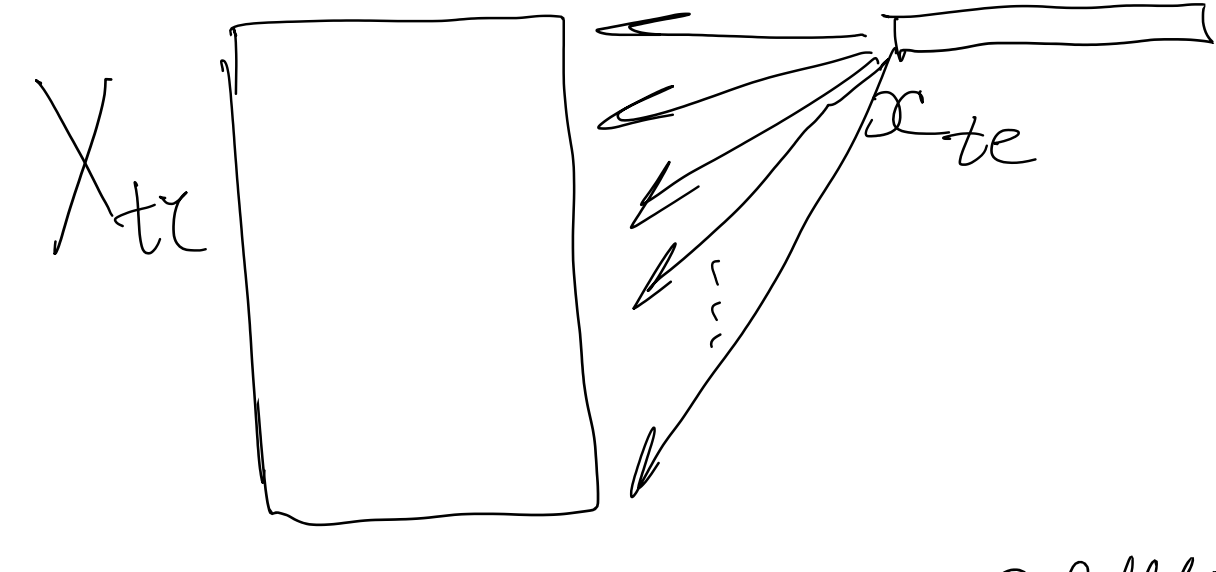
Обучение с учителем $x \in \mathbb{R}^d, y$
регрессия, классификация
 $y \in \mathbb{R}$ $y \in \{1, \dots, K\}$

- методы - kNN $x = \{x_1, x_2, x_3, \dots, x_d\}$
- метрики расстояния
- подготовка данных

Дано:
 X  Y 
 $X \in \mathbb{R}^{l \times d}$ $Y \in \mathbb{R}^l$ - регрессия
 $Y \in \{1, \dots, K\}$ - классификация
обучающие данные

Найти: алгоритм a
 $a(x) \approx y$

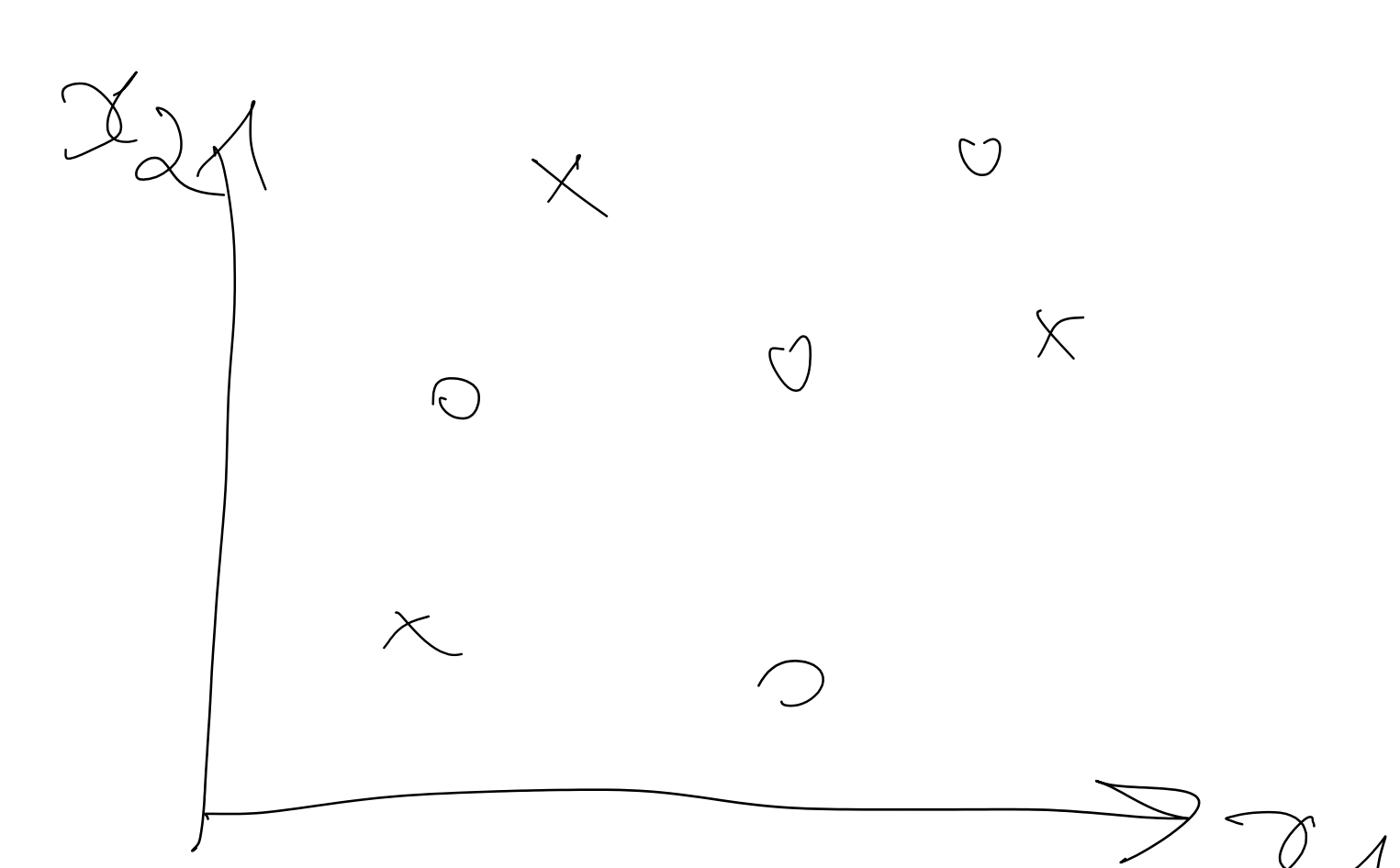
Метод k ближайших соседей (kNN):



1. Измерить расстояния $\rho(x_{te}, x_i)$ $i=1, \dots, l$
 2. Выбрав k ближайших соседей
 3. Присвоение = усреднение / по соседям / по соседям
- предск. $\frac{\text{по соседям}}{\text{по соседям}}$ $\frac{\text{по соседям}}{\text{по соседям}}$

Обучение: запомнить обучающие данные

Функции расстояния:
евклидова: $\rho(x, z) = \sqrt{\sum_{i=1}^d (x_i - z_i)^2}$
манхэттенская: $\rho(x, z) = \sum_{i=1}^d |x_i - z_i|$



Задача 1 $x_{te} = (2, -1)$ $x_1 = (1, -1)$
Метод: $\rho(x_{te}, x_1) = |2-1| + |-1-(-1)| = 1$

$$\rho(x_{te}, x_2) = |2-2| + |-1-(-2)| = 1$$

$$\rho(x_{te}, x_3) = |2-3| + |-1-(-2)| = 1$$

$$\rho(x_{te}, x_4) = |2-1| + |-1-(-1)| = 1$$

$$\rho(x_{te}, x_5) = |2-2| + |-1-(-2)| = 1$$

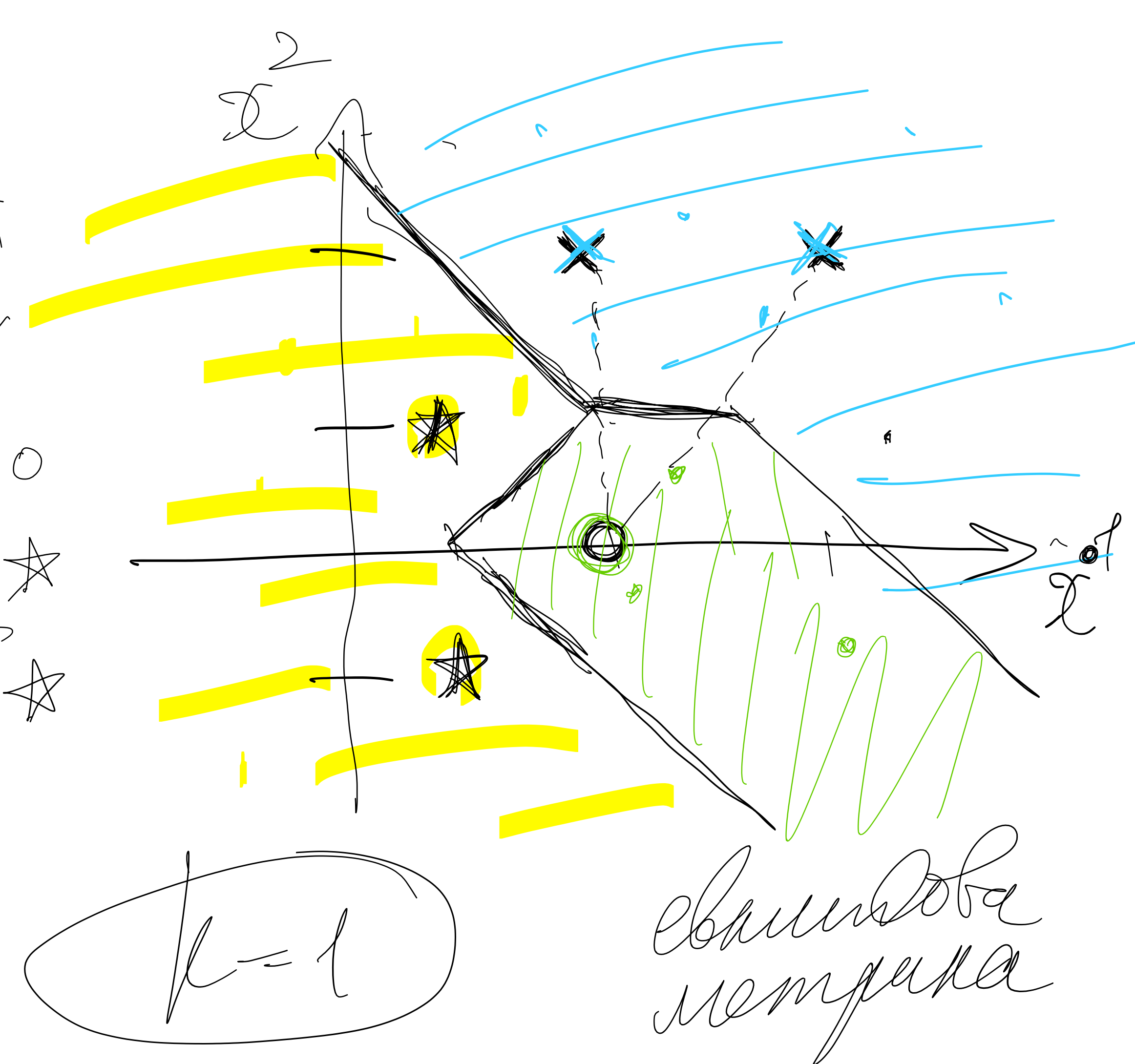
Метод k=3 ближайшее по координатам
 $i = 1, 4, 5$

Метод 3: $y_1 = 1$ $y_4 = 3$ $y_5 = 3$

выбравшие
 $\rightarrow a(x_{te}) = 3$

Задача 2

x^1	x^2	y	
2	2	1	x
3	2	1	x
2	0	2	o
1	-1	3	*
1	1	3	*



$k=1$

евклидова метрика

Задача 3

1	-1	3.5	V
2	2	2.3	
3	2	1.7	V
1	0	-0.9	V
2	-2	0.4	V

Ближайшие соседи;

$i = 1, 4, 5$

$$a(x) = \frac{3.5 + 0.4 + 0.1}{3} = \frac{4.0}{3} \approx 1.33$$

Задача регрессии
yov

Эвристика — идея, которая, кажется, будет работать, но теоретически не гарантируется
+ еще хорошо работает на практике

Гиперпараметры алгоритма

Для kNN: — k — число соседей



— функции расстояния

— процедура выбора весов