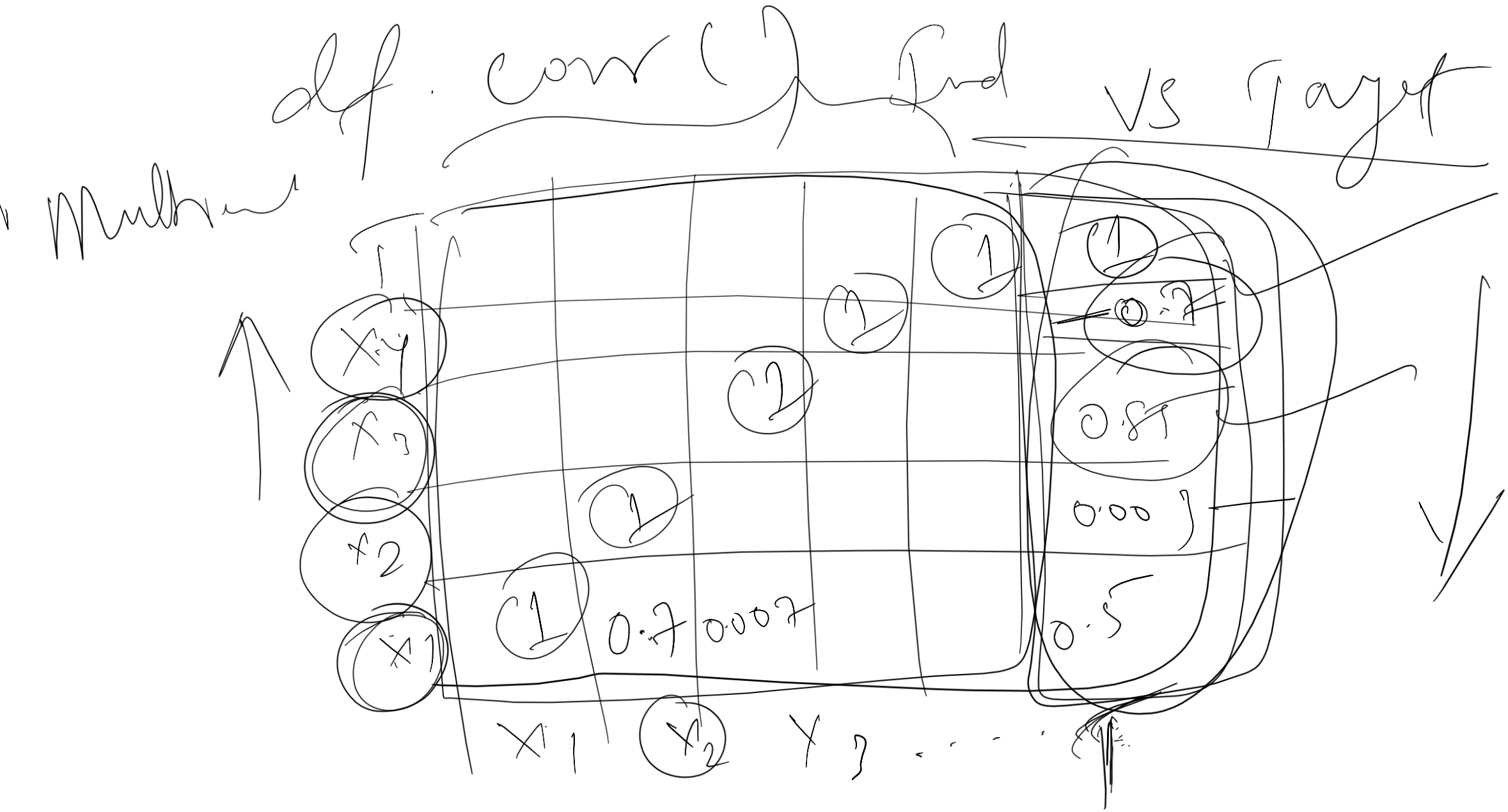
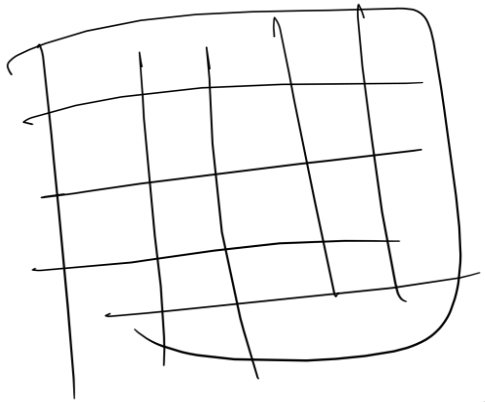
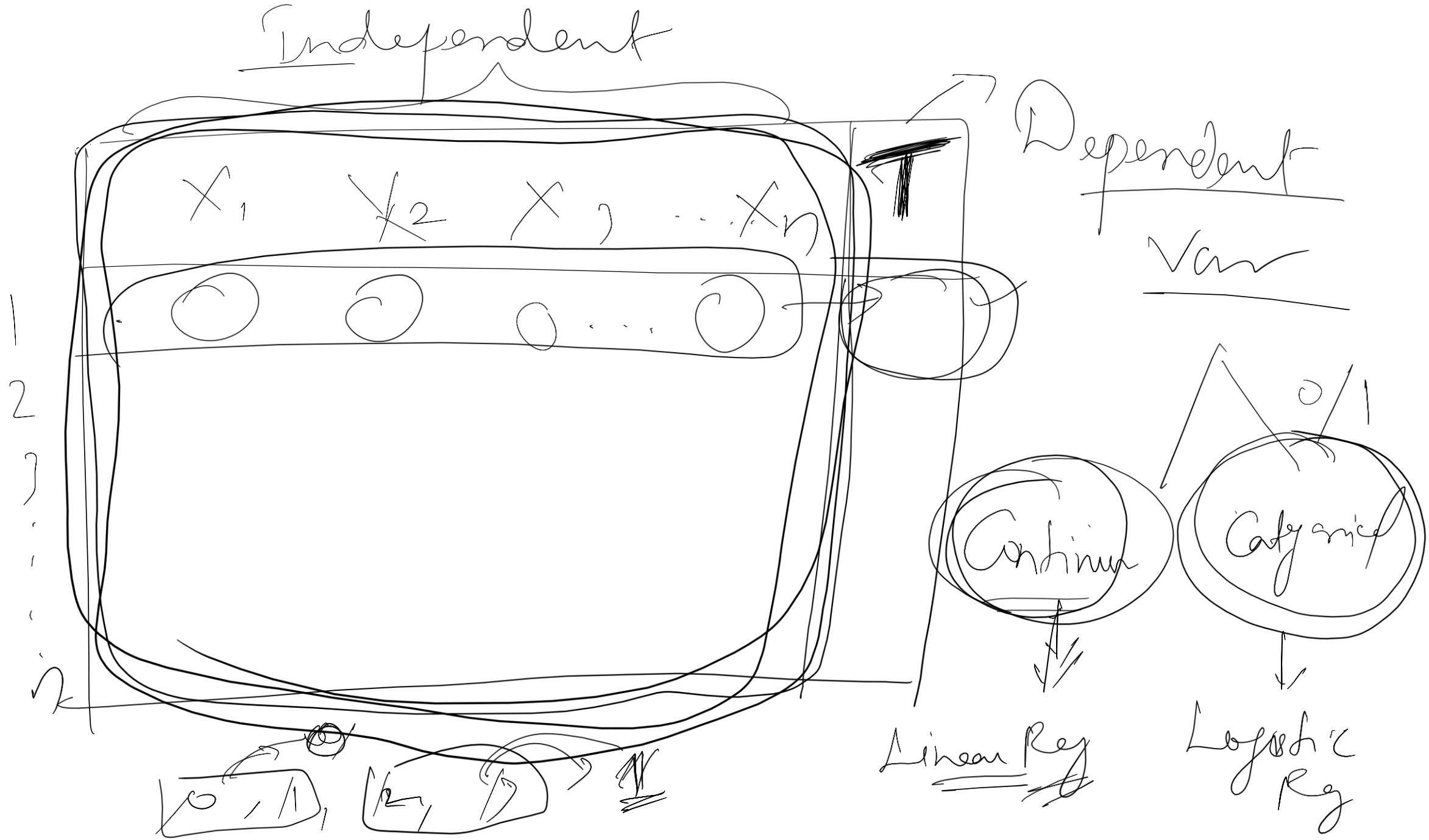

300
Hist
Freq
200
100
0
0
1
1 var
Layer
n = 500
200 = 0
300 = 1


df.corr()
Vs Target
1
1
1
1
0.7
0.84
0.001
0.5
0.7
0.0007
X1
X2
X3
X4
X1
X2
X3

EDA → 

→ no null
→ int/flts

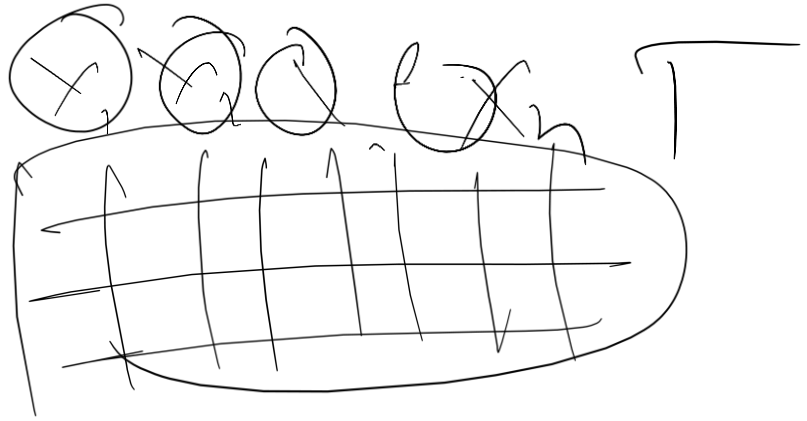
```
model = LinearRegression()
model.fit(X_train, y_train)
model.predict(X_test)
```


Independent
X_1 X_2 X_3 ... X_n
Dependent Var
T
1
2
3
n
0, 1
Continuous
Categorical
0/1
Linear Reg
Logistic Reg

Supervised	Unsupervised
Target	No target

- Linear Regression
- Logistic Regression
- KNN

Linear Regression :



$x_1, x_2 .. x_n$ → Coeff

β_0 → Intercept

$$mx + b$$

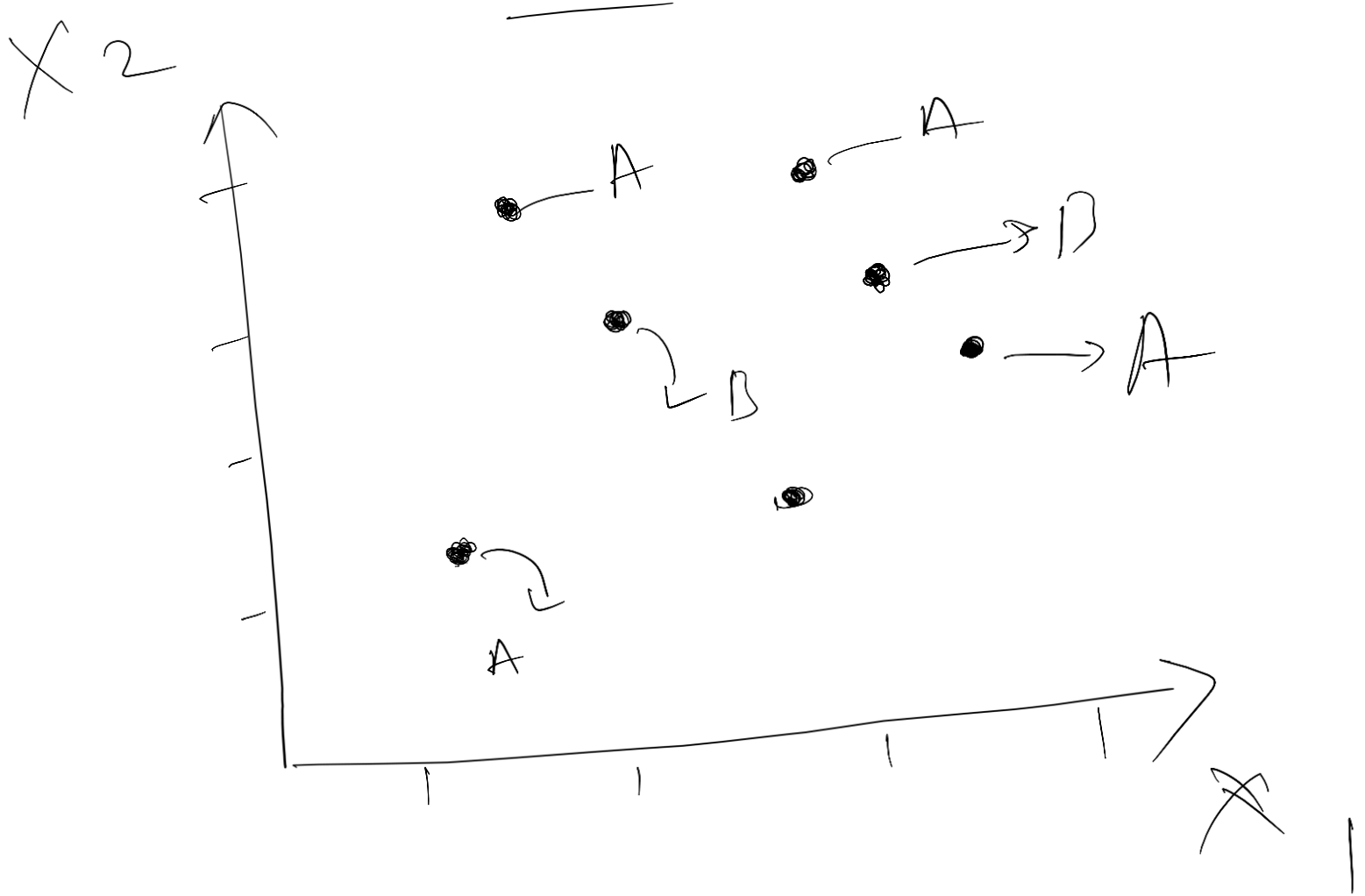
$$\beta_0 + \beta_1 x$$

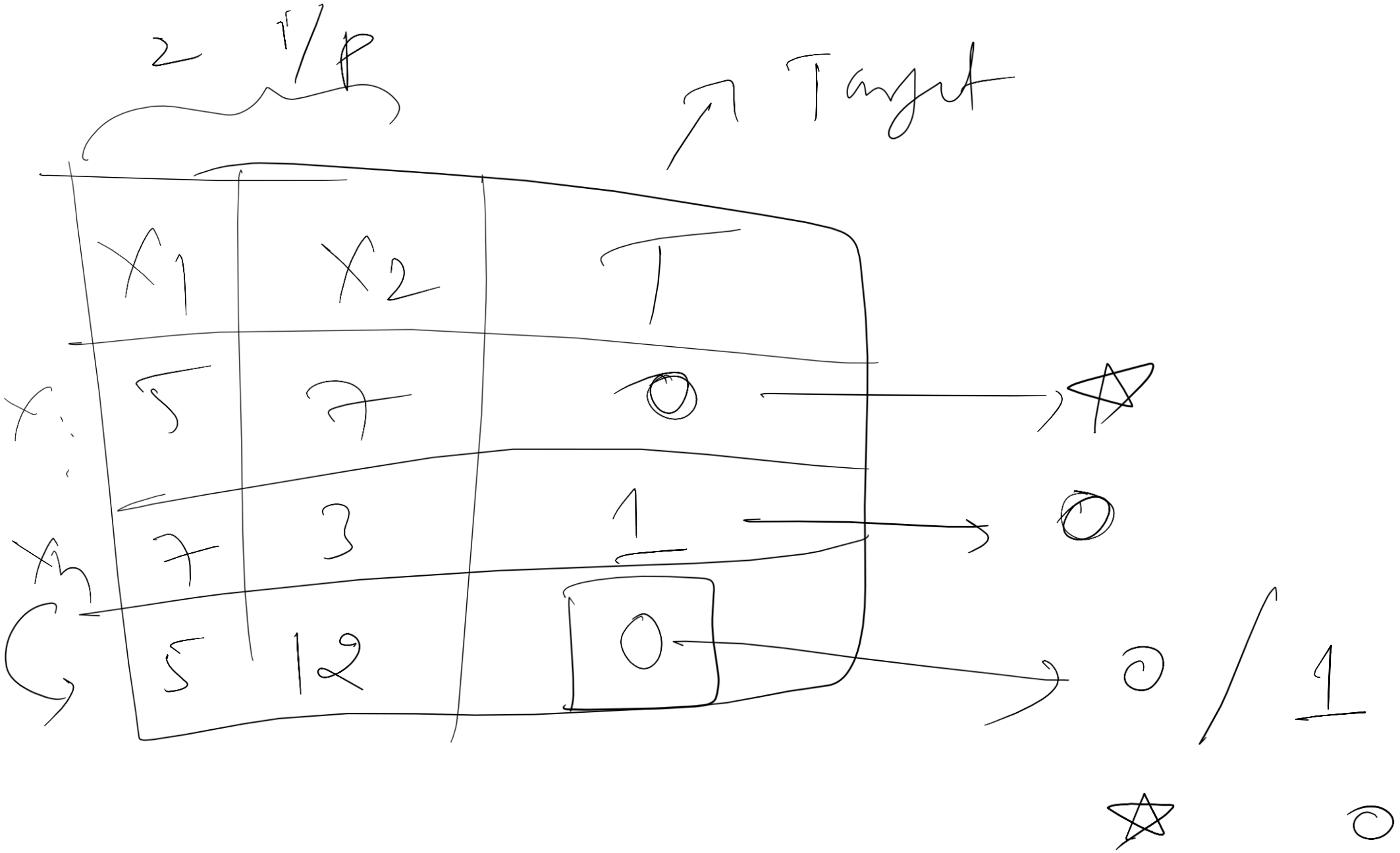
KNN (K - Nearest Neighbor)

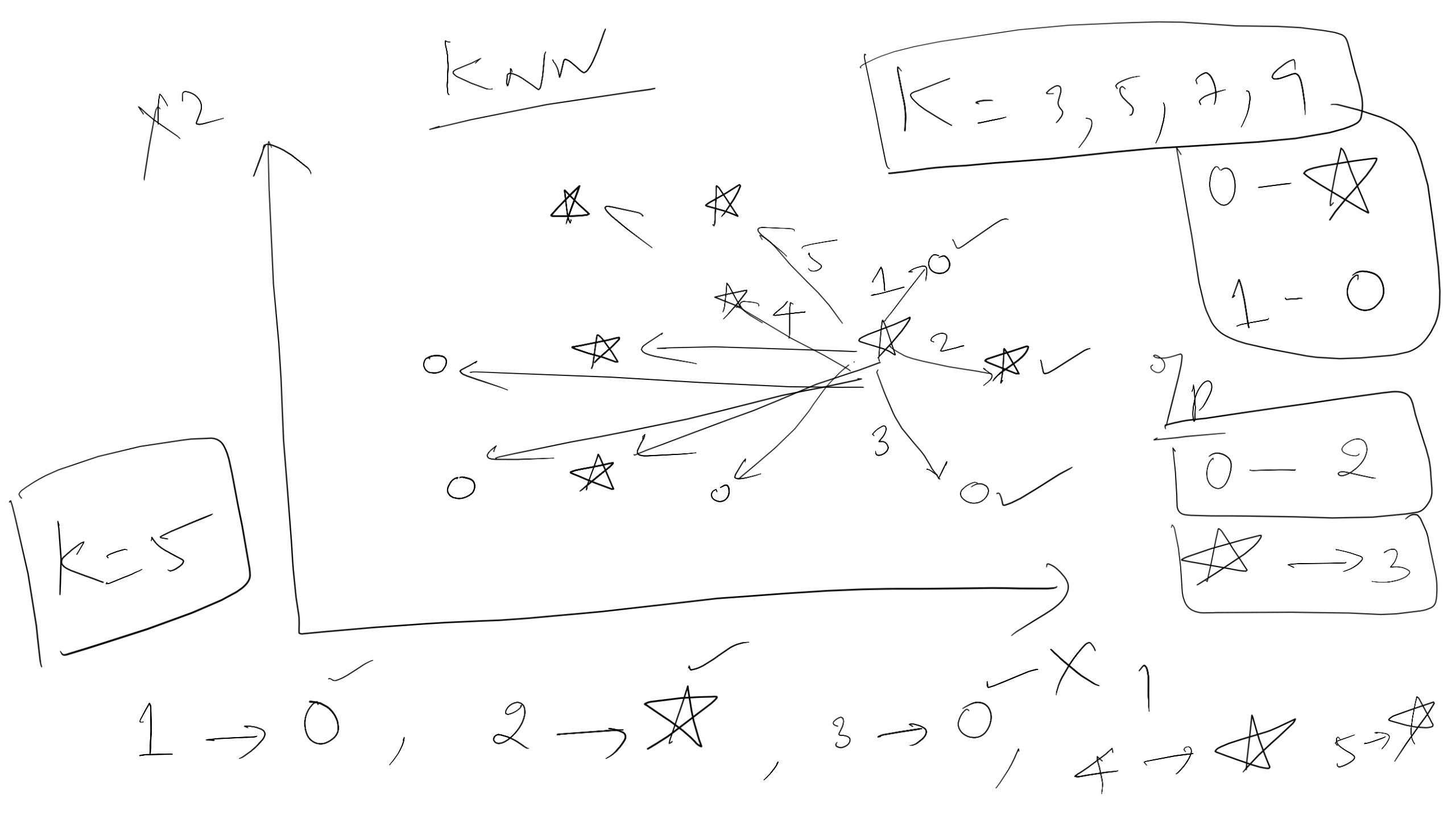
↳ Super-simple Supervised Learning algo

↳ Lazy learning algorithm

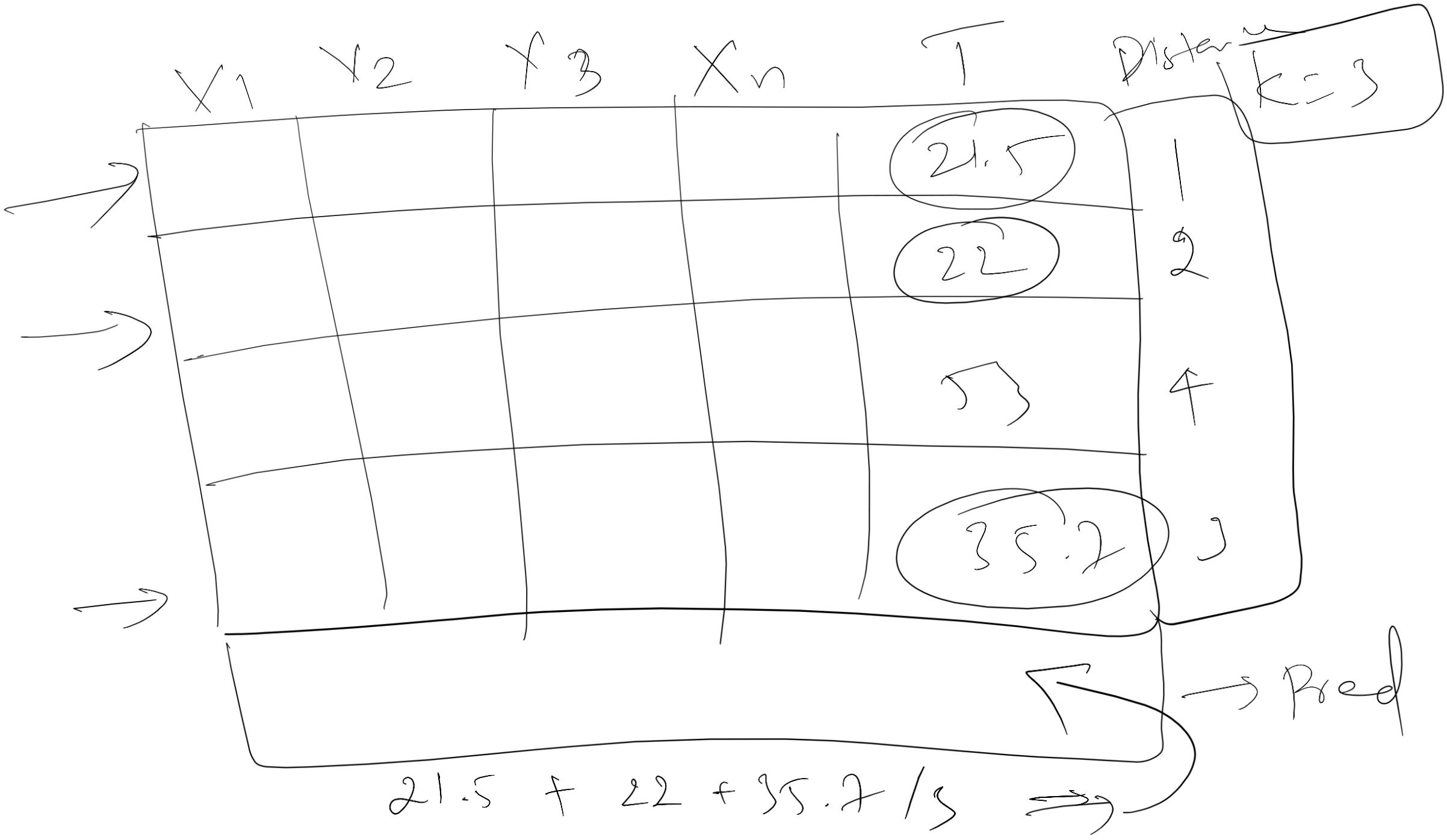
2.9




2 I/P
Target
x_1
x_2
T
5
7
0
x_i
7
3
1
x_m
5
12
0
0 / 1


KNN
x2
K = 3, 5, 7, 9
0 - ☆
1 - 0
Zp
0 - 2
☆ → 3
K - 5
1
2
3
4
5
1 → 0 , 2 → ☆ , 3 → 0 , 4 → ☆ 5 → ☆


X1
X2
X3
T
Distance
3 4 3 0 70
7 3 2 1 20
5 2 3 1 50
2 2 1 0 100
1 5 7 1
k=3
0 → 1
1 → 2

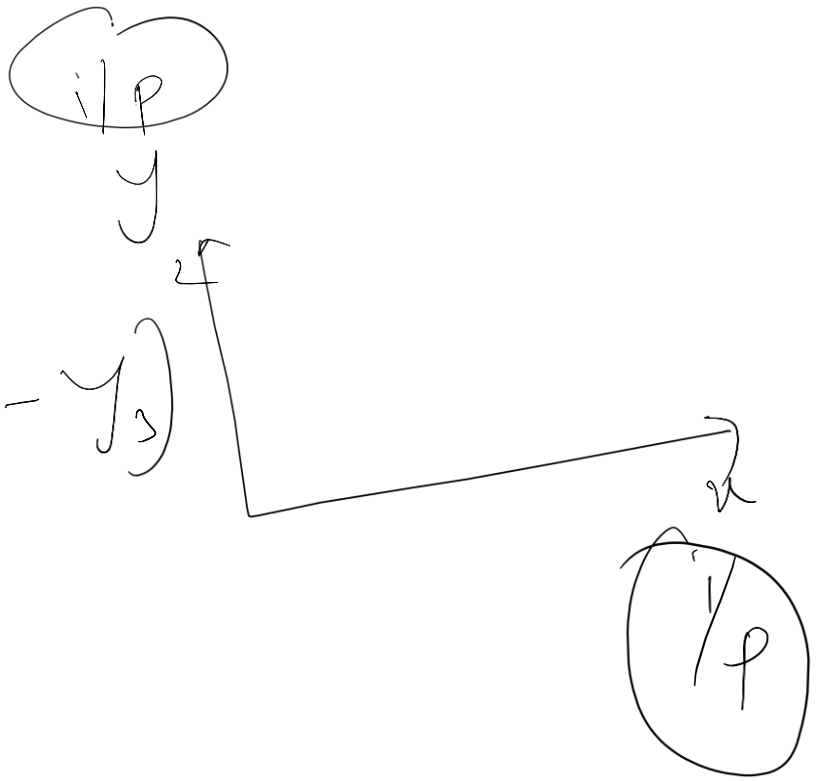

X1
X2
X3
Xn
T
Distance
K=3
21.5
22
35.7
1
2
4
3
Pred
21.5 + 22 + 35.7 / 3

Classification → Mode

Regression → Mean

Distance

$$\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}$$



$$\sqrt{\sum_{i=1}^{3} (x_i - y_i)^2}$$

$i \rightarrow$ Record

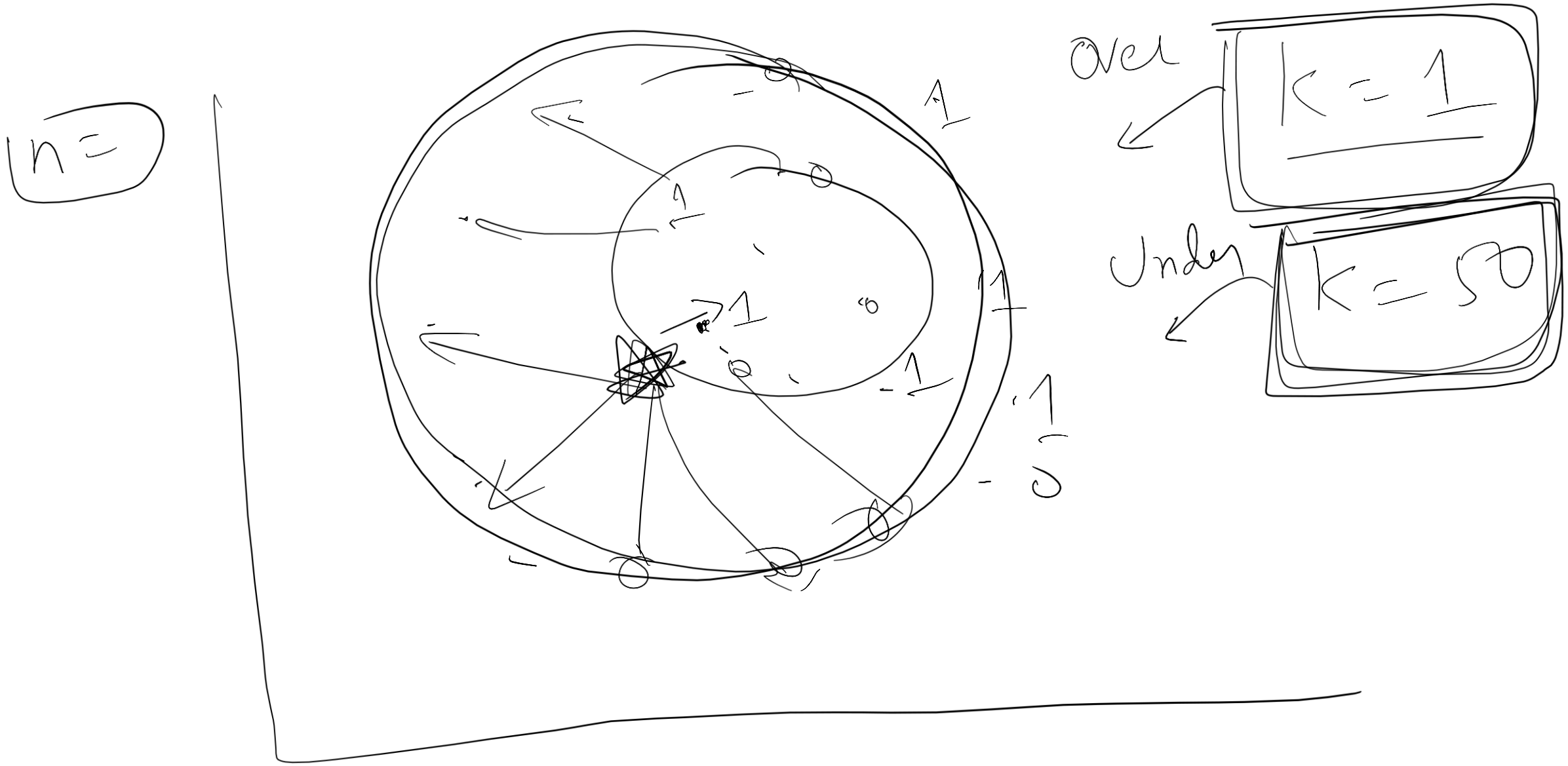
$$\left(\left(x_1 - y_1\right)^2\right)^{1/2}$$

$$\sqrt{\left(x_1 - y_1\right)^2}$$

$$-9 = 2$$

$$- = 1$$

$$- = 7$$


n=
over
k=1
Under
k=50

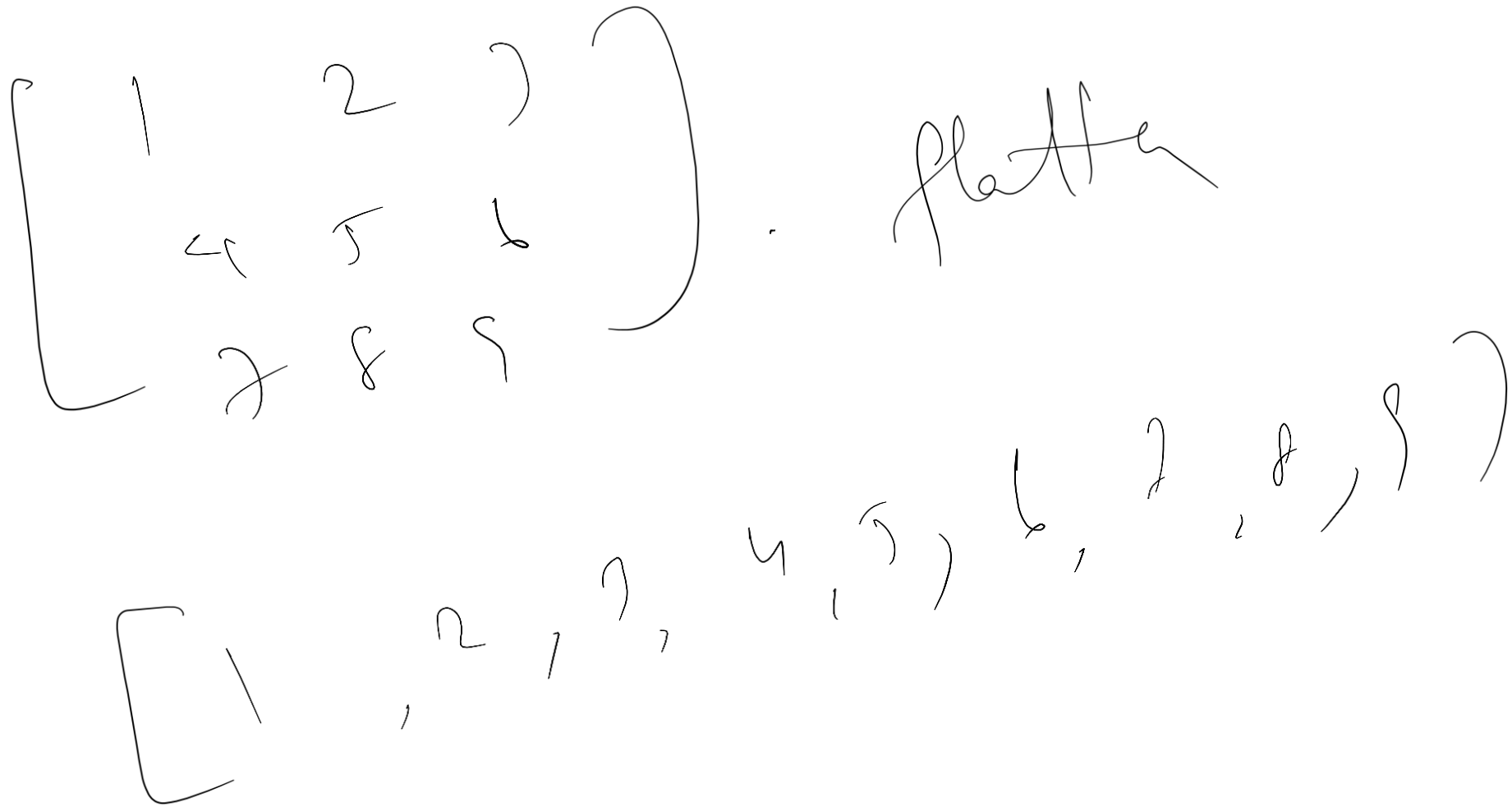
$$x_1 \quad x_2 \quad x_3 \quad \cdots \quad x_{1,500}$$

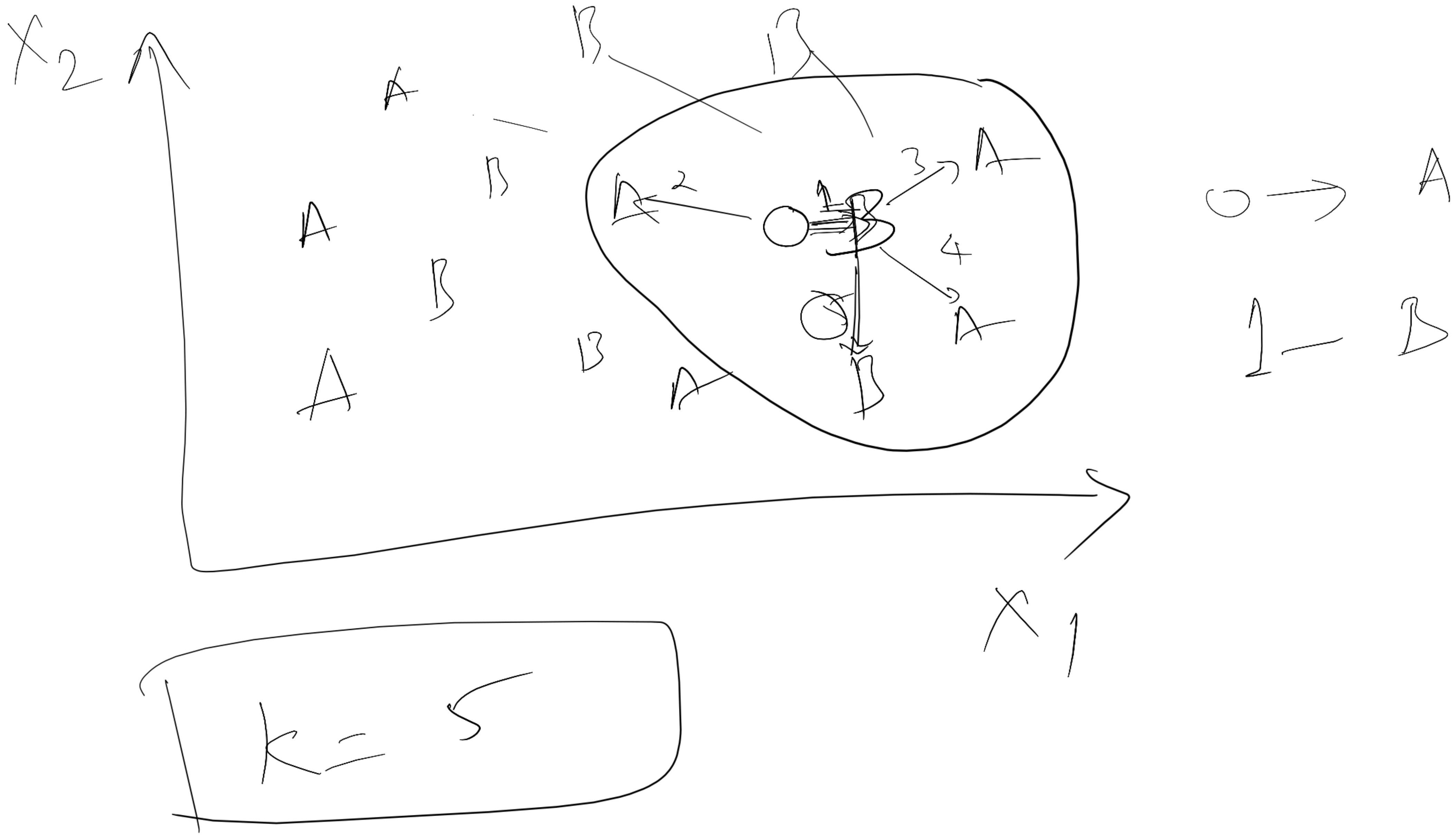
$$(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + \cdots$$

$$\left(x_{1,000} - y_{1,000} \right)^2$$

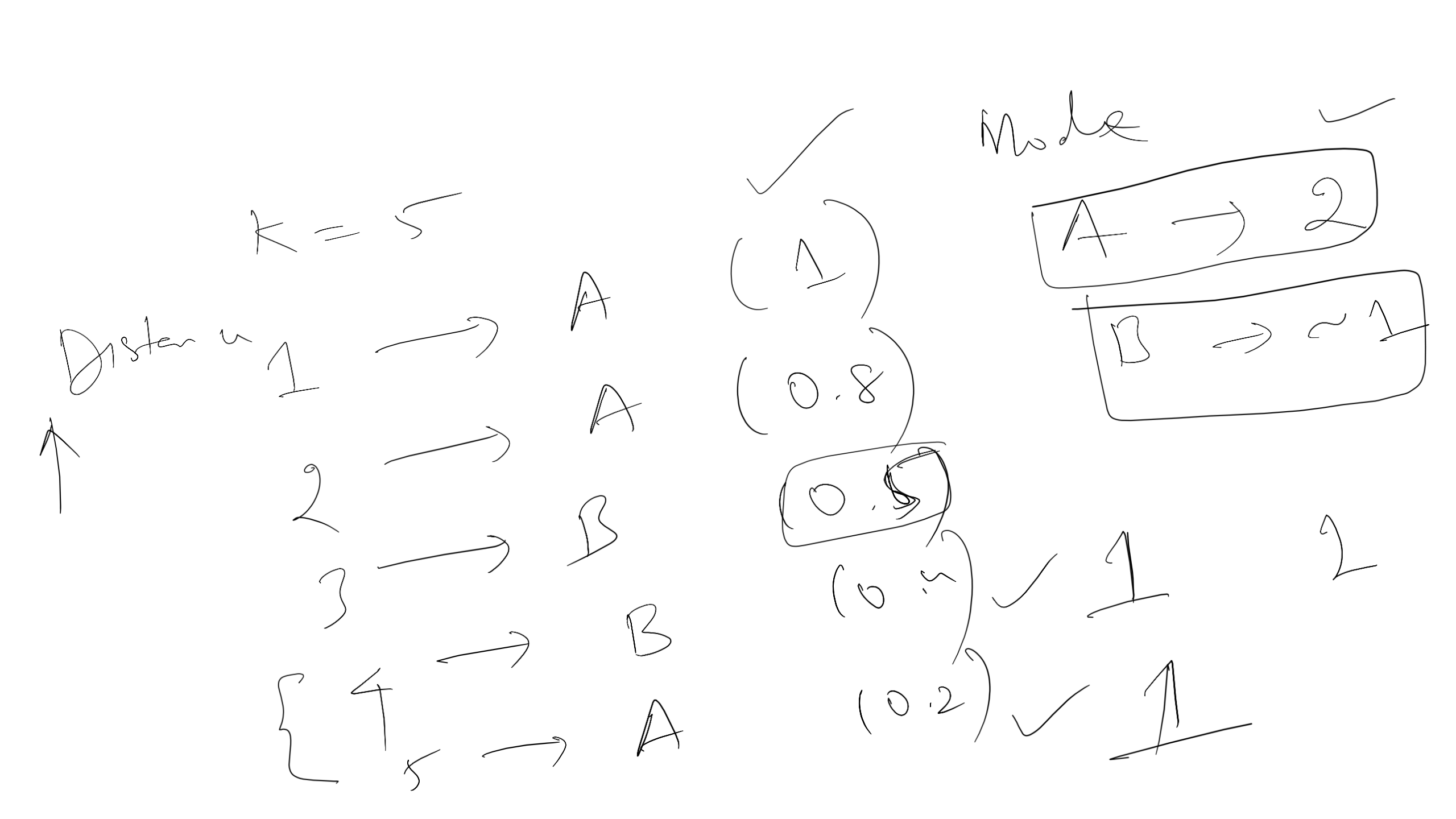
Let's get back

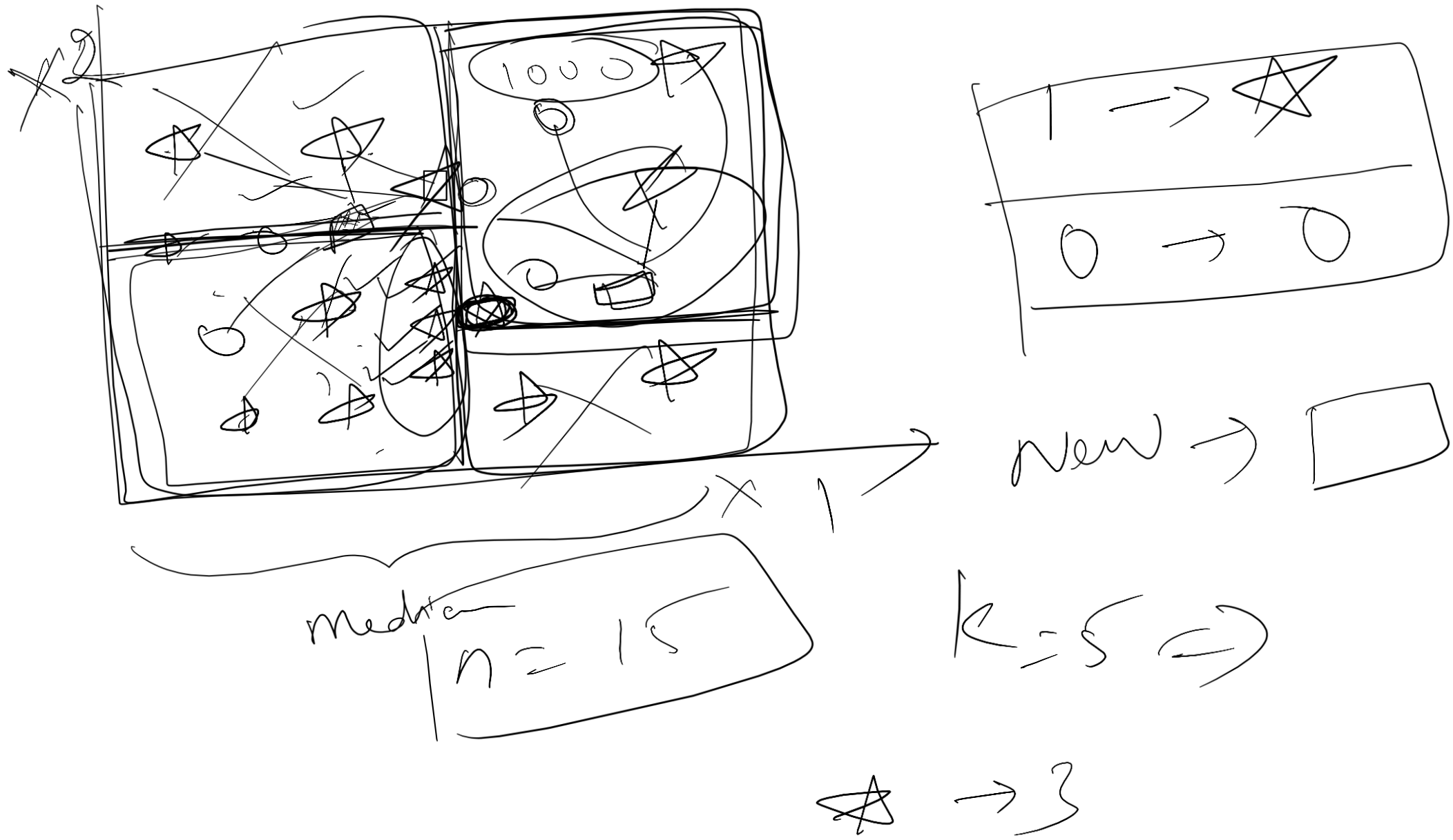
7:55 pm. IST

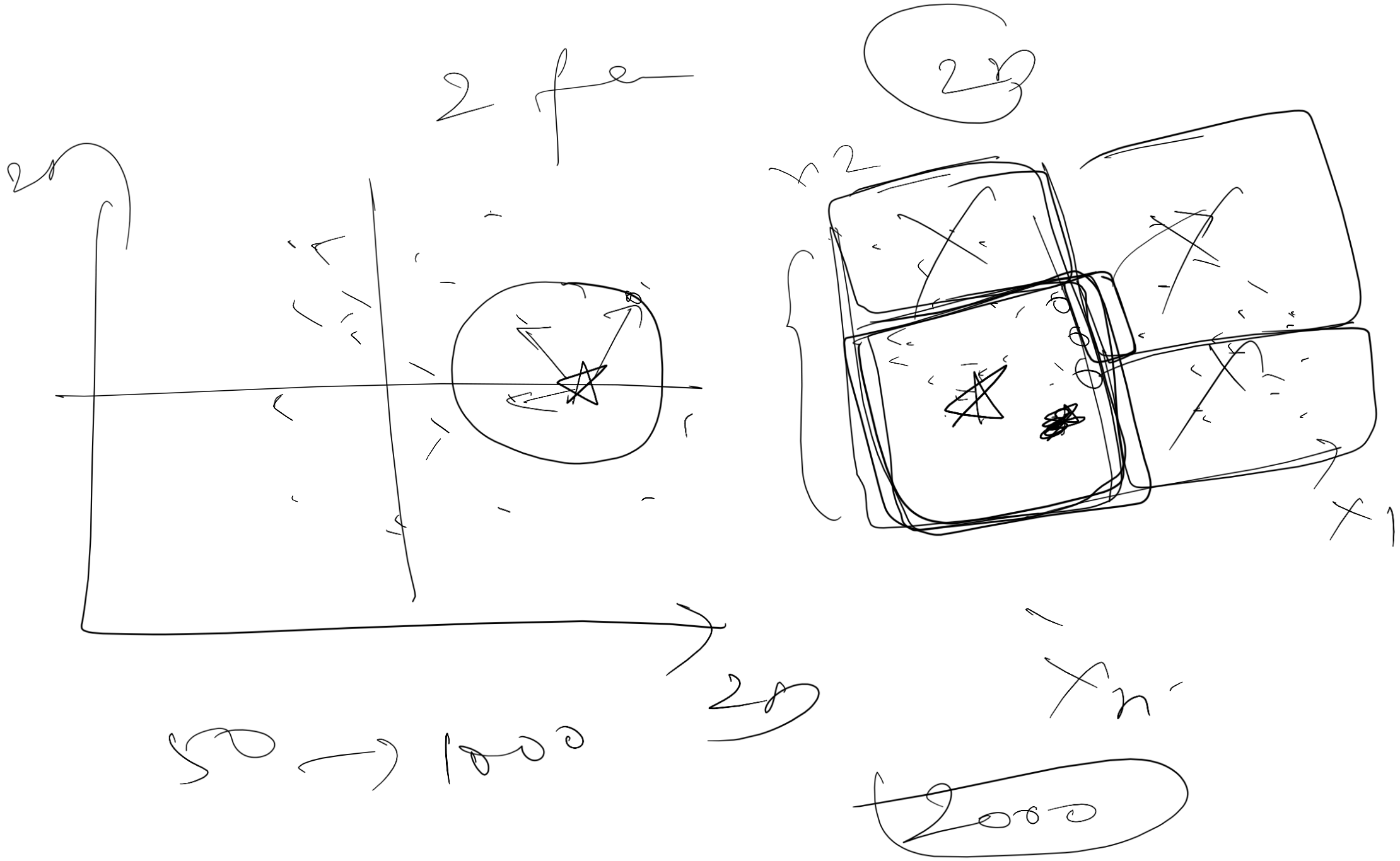


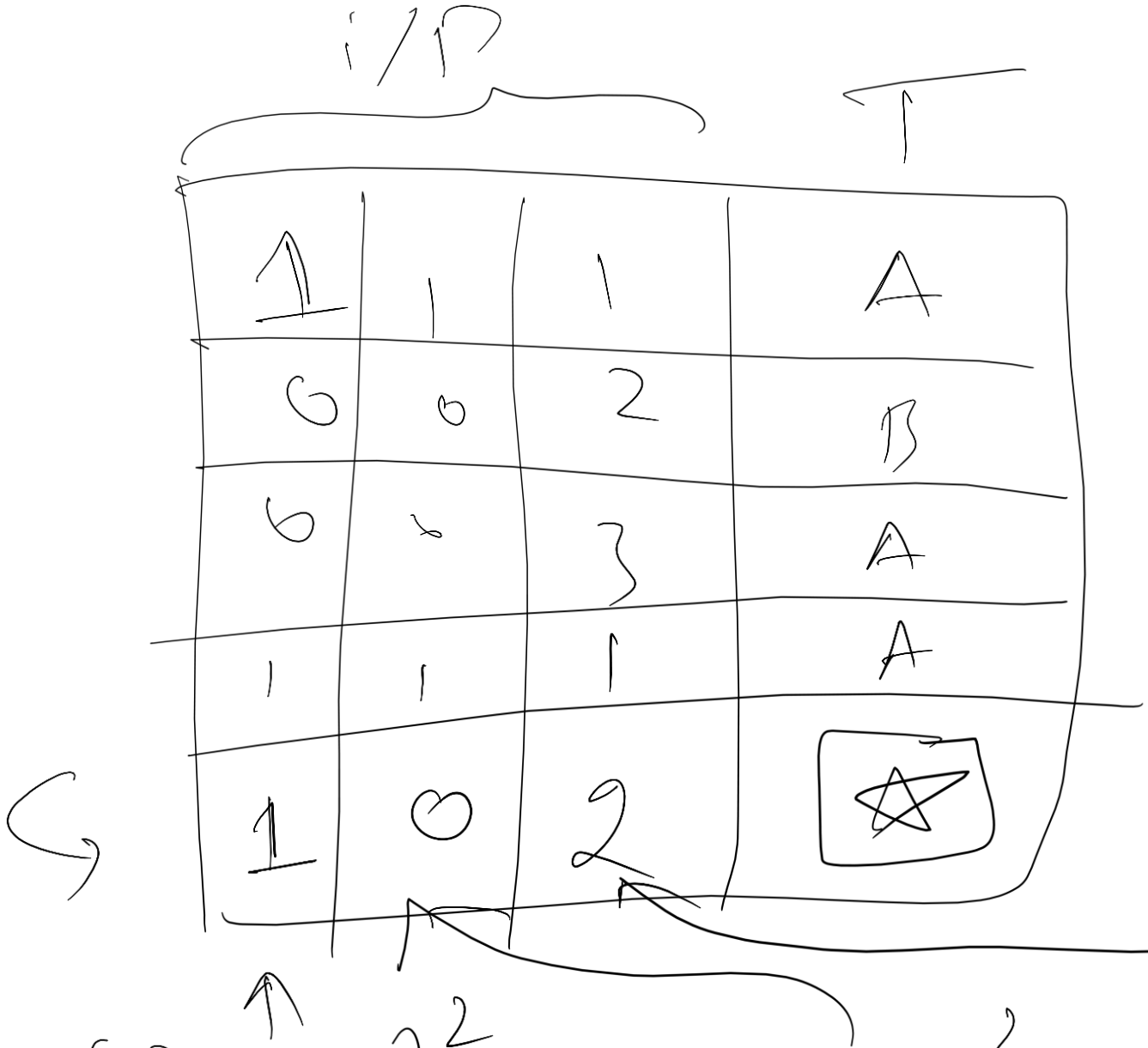


X2
X1
k = 5


K = 5
Distance
1 → A
2 → A
3 → B
4 → B
5 → A
(1)
(0.8)
(0.5)
(0.4)
(0.2)
Mode
A → 2
B → 1
1
1
1


x2
100
1 → ☆
0 → 0
new → □
x 1
median
n = 15
k = 5 ⇒
☆ → 3





Binary

Multi-class

A → 0

B → 1

$(x_1 - y_1)^2 + (x_1 - y_1)^2 + (x_1 - y_1)^2$

$6 + 1 + 1 = 2$

⇒ knn

Screening Round : (1) MAANG

E/m/H → Online Assignment (mcq) → Python / D.S.

→ Coder Pad (Live-coding)

Leetcode → sql, pandas (Python)

Technical Assessment : (V.O) 4 ⇒ Sa

(1-2) (2-4)

M.L