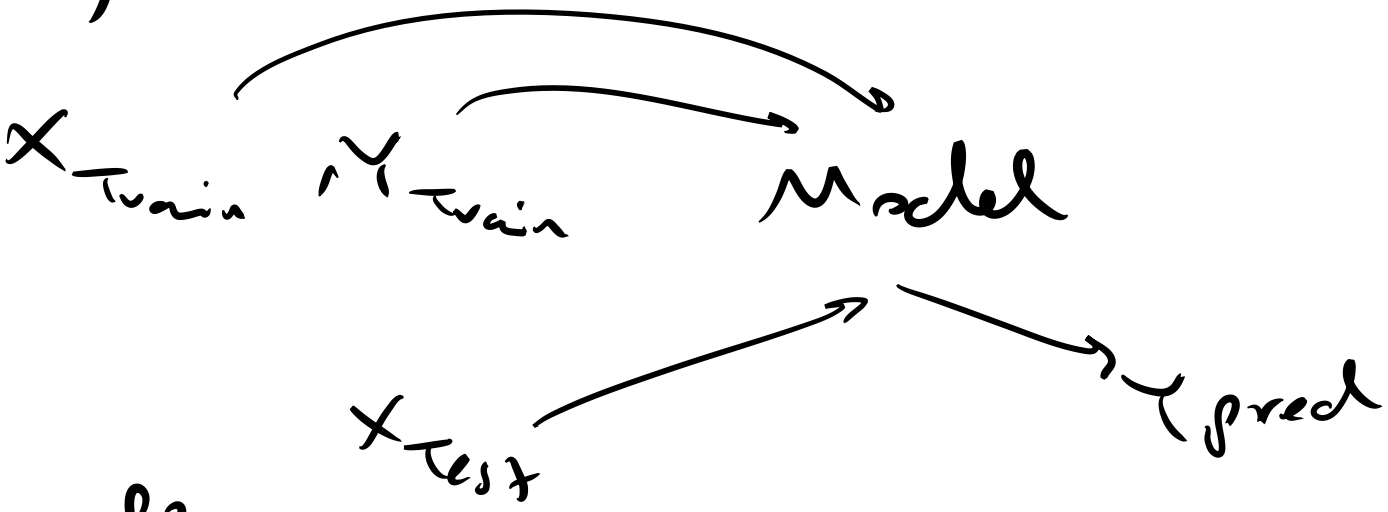
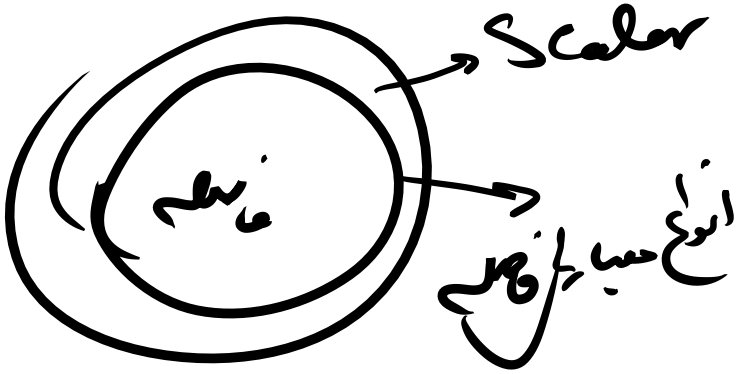
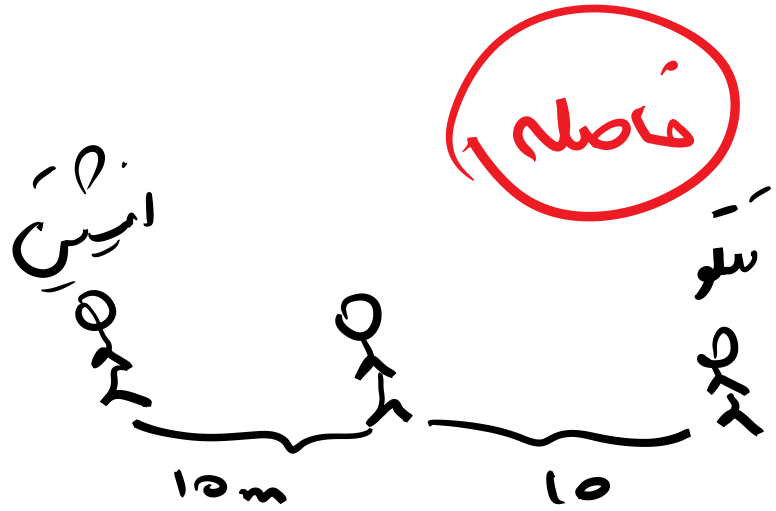
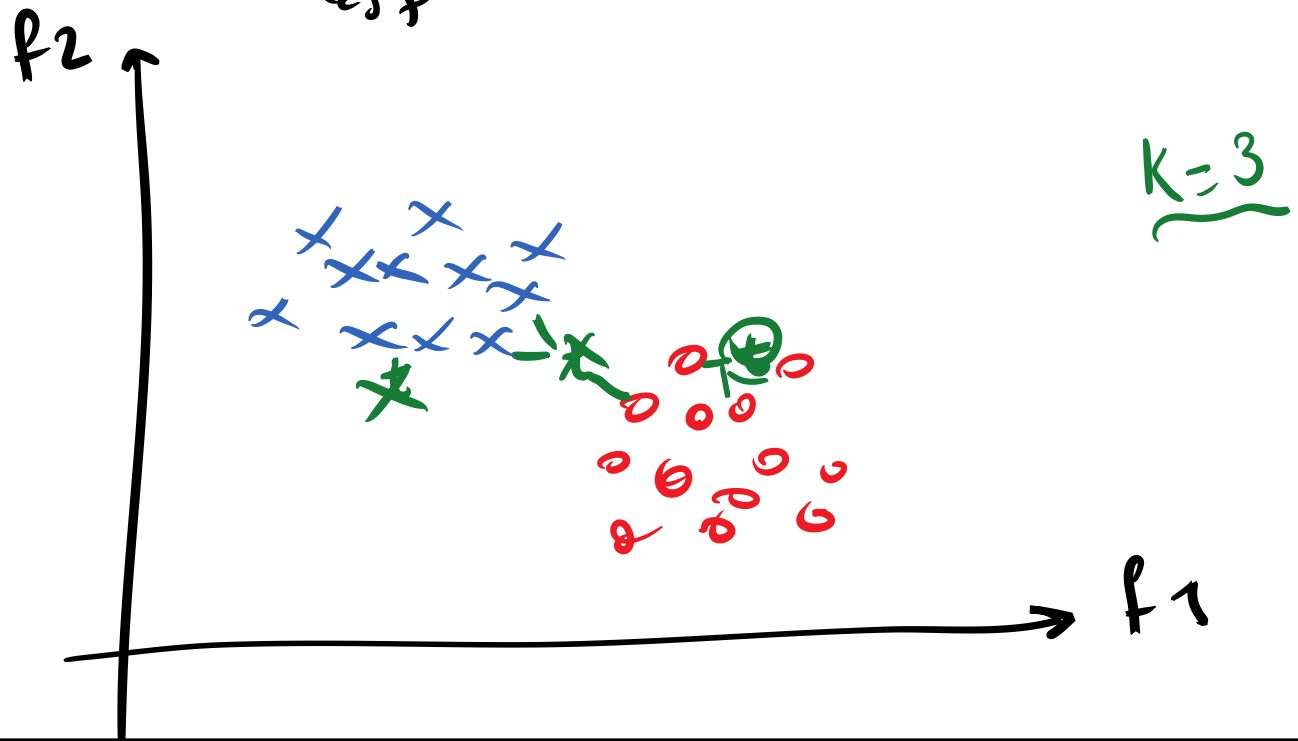
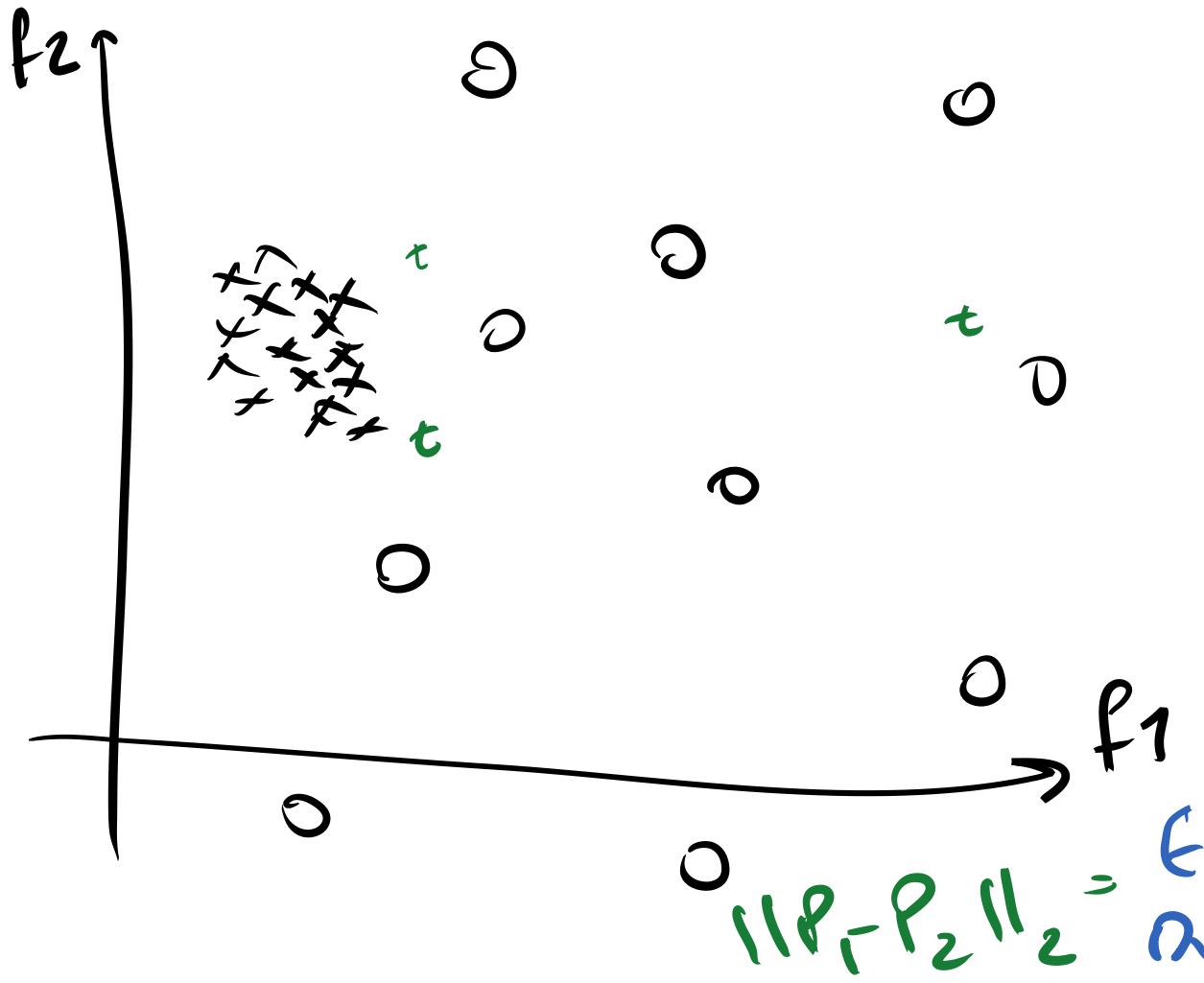
7) KNN


X_train
Y_train
Model
X_test
Y_pred


Scaler
نرمال
انواع مسافت فاصله


فاصله
15m
15


f2
f1
K=3



$\|P_1 - P_2\|_2$ = Euclidian Distance

f1 f2 f1 f2

(2, 5) (3, 8)

$$\sqrt{(2-3)^2 + (5-8)^2}$$

P_1 P_2

$$\sqrt{(P_1 - P_2)^T (P_1 - P_2)}$$

1×2 2×1

$$Mahalanobis\_Distance_{(P_1,P_2)} = \sqrt{(P_1 - P_2)^T \Sigma^{-1} (P_1 - P_2)}$$

$(P_1 - P_2)$: $1 \times f$ — Σ^{-1}: $f \times f$ — $(P_1 - P_2)$: $f \times 1$

ماتریس کوواریانس Σ

$$\Sigma = \begin{bmatrix} Var(x_1) & Cov_{1,2} & Cov_{1,3} \\ Cov_{2,1} & Var(x_2) & Cov_{2,3} \\ Cov_{3,1} & Cov_{3,2} & Var(x_3) \end{bmatrix}_{3 \times 3}$$

$$Cov(x,y) = \frac{1}{n} \sum (x_i - \mu_x)(y_i - \mu_y)$$

$$\sigma^2 = Var(x) = \frac{1}{n} \sum (x_i - \mu)^2$$

$N(\mu, \sigma)$ $N(\mu, \Sigma)$

Manhattan_distance $= \sum_{i=1}^{n} |x_i - y_i| = \|x - y\|_1$

(x, y)

Chebychef_dist $= \max_{i=1}^{n} (x_i - y_i) = \|x - y\|_\infty$

(x, y)

Minkowsky_dist $= \left(\sum |x_i - y_i|^p \right)^{1/p}$

(x, y)

$p = 1$

$p = 2$ — اقلیدسی

$p = \infty$

$P_1 = (4, 7, 3)$

$P_2 = (6, 2, 8)$

Manhattan Dist $(P_1, P_2) = |4-6| + |7-2| + |3-8|$

$= 12$

$\|P_1 - P_2\|_1$

Euclidean_Distance $(P_1, P_2) = \sqrt{(4-6)^2 + (7-2)^2 + (3-8)^2} = \sqrt{54}$

$\|P_1 - P_2\|_2$

Chebychev $(P_1, P_2) = 5$

$\|P_1 - P_2\|_\infty$

Categorical features

$P_1 = (\text{Male}, \text{'Math'})$ → $(0, 0)$

$P_2 = (\text{female}, \text{'Computer'})$ → $(1, 1)$

$P_3 = (\text{Male}, \text{'Art'})$ → $(0, 2)$

$P_4 = (\text{female}, \text{'physics'})$ → $(1, 3)$

$\|P_1 - P_2\|_2 = \sqrt{(0-1)^2 + (0-1)^2}$

$\|P_1 - P_4\|_2 = \sqrt{(0-1)^2 + (0-3)^2}$

hamming(P_1, P_2)

$= 1 + 1$

hamming$(P_1, P_3) = 0 + 1$

هر کسی که

ترتیب هم

کودک	نوجوان	جوان	سن
1	2	3	۴

نوزاد		اردیبهشت	فروردین
2		1	0

$P_1 = (\text{Female}, \text{'Math'}, 170, 75)$ → $(0, 3, 170, 75)$

$P_2 = (\text{Male}, \text{'Computer'}, 160, 90)$ → $(1, 2, 160, 90)$

$P_3 = (\text{Female}, \text{'Art'}, 190, 60)$ → $(0, 1, 190, 60)$

$P_4 = (\text{Male}, \text{'Art'}, 160, 120)$ → $(1, 1, 160, 120)$

range

$$\text{Distance}(P_1, P_3) = \sqrt{(0-0)^2 + 1^2 + |170-190|^2 + |75-60|^2}$$

Scaler

- MinMaxScaler → (0, 1)
- Standard Scaler

$$\frac{x_i - \min}{\max - \min} \sim U(0,1)$$

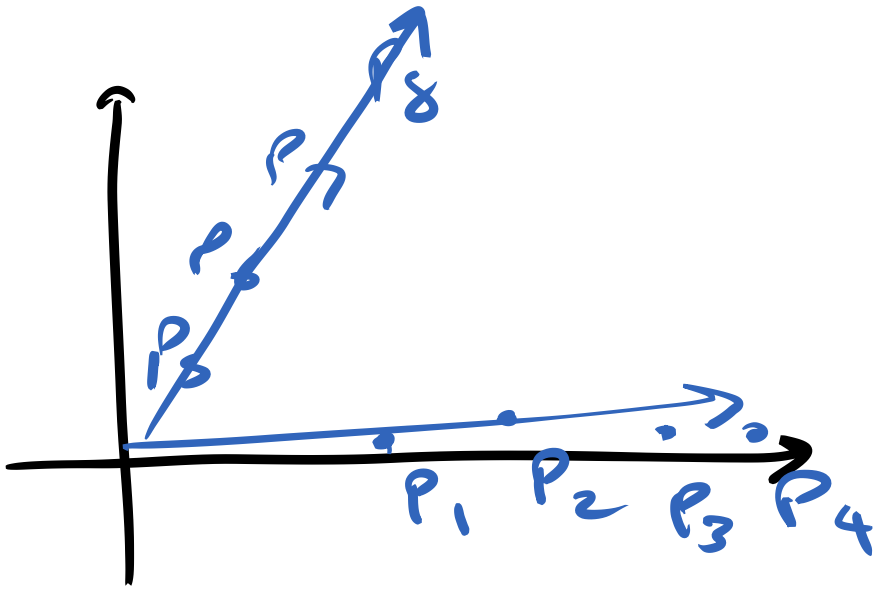
$N(10, 5)$

$$z_i = \frac{x_i - \mu}{\delta} \sim N(0,1)$$

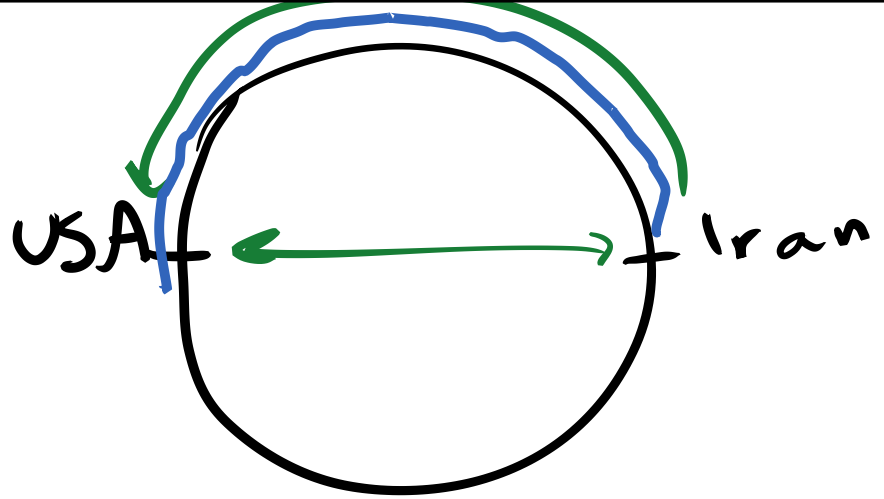
$(-3, 3)$

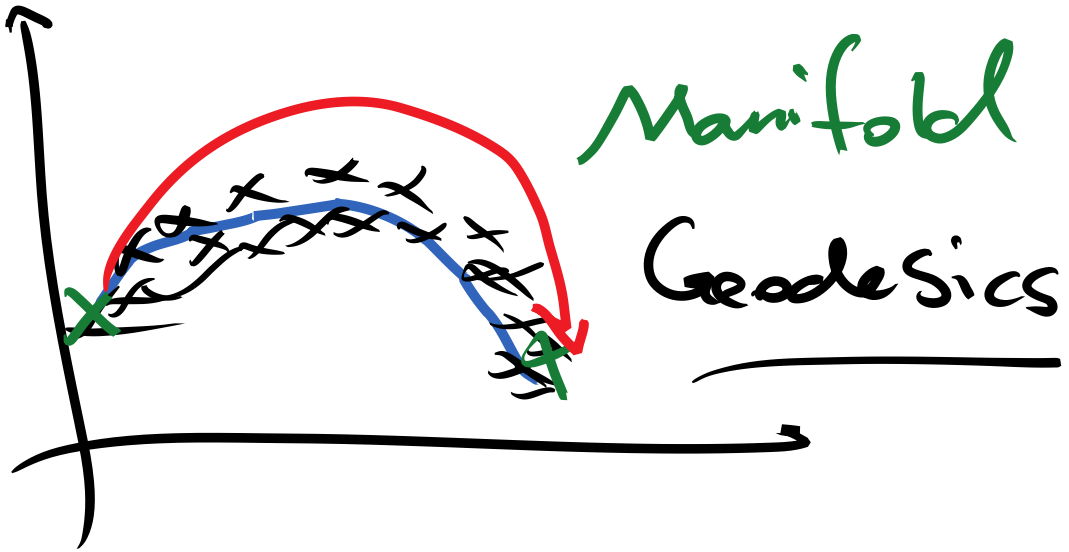
Cosine Similarity $|x \cdot y| = |x| \cdot |y| \cos\theta$

ضرب داخلی دو بردار



$$\cos\gamma = \frac{|x \cdot y|}{|x||y|}$$





Manifold

Geodesics

Metric learning

$$\sqrt{(p_1 - p_2)^T \Sigma^{-1} (p_1 - p_2)}$$

$$M$$