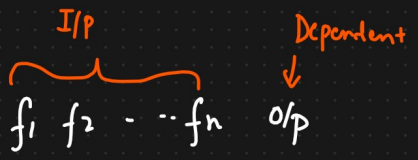
Unsupervised Machine Learning

Clustering Algorithms

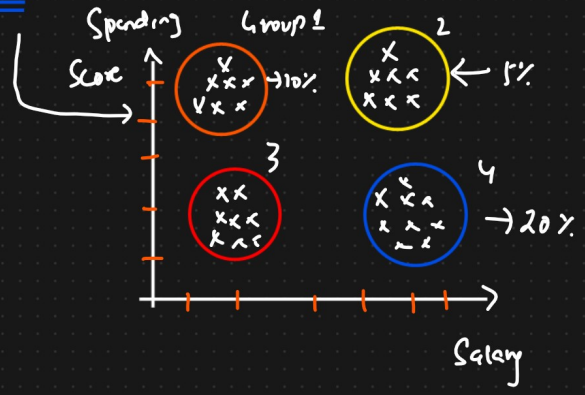
1. K Means Clustering
2. Hicranichal Clustering.
3. DBScan Clustering.



Supervised ML

f1	f2	f3	- -	fn
-	-	-		-
-	-	-		-

Ecommerce → Iphone 20 → 1.8 rakhs

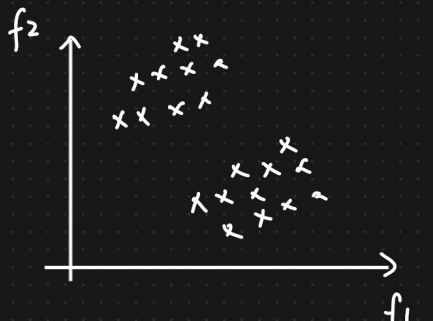


Dataset

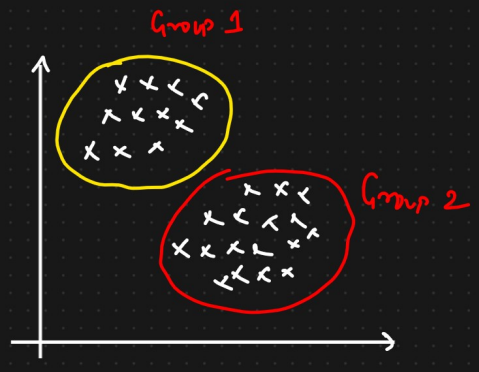
Salary	Spending Score (1-10)
-	-
-	-
-	-

1. K Means Clustering

Geometric Intuition

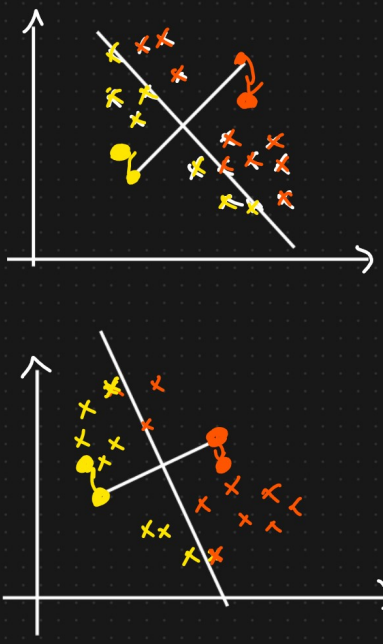


⇒ K Means



⇒ K Means

K Means Mathematical Intuition



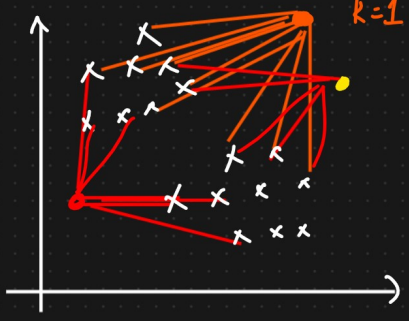
Steps: $k=2$

1. Initialize some K → centroids
2. Points that are nearest to the centroid → Group
3. Move the centroids → Mean

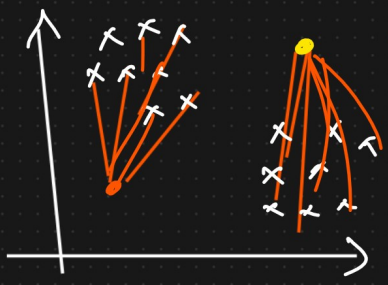
How do we select the K Value ?

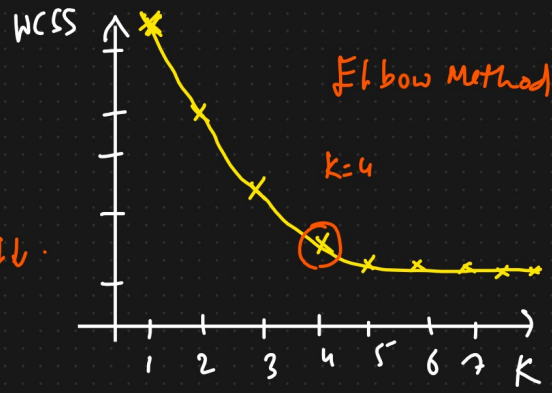
WCSS = Within Cluster Sum of Squares

Initialize K = 1 to 20

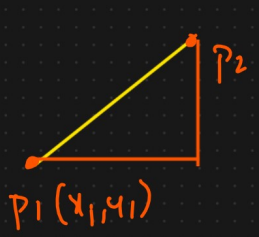


$$WCSS = \sum_{i=1}^{K} (\text{Distance between point to the nearest centroid})^2$$





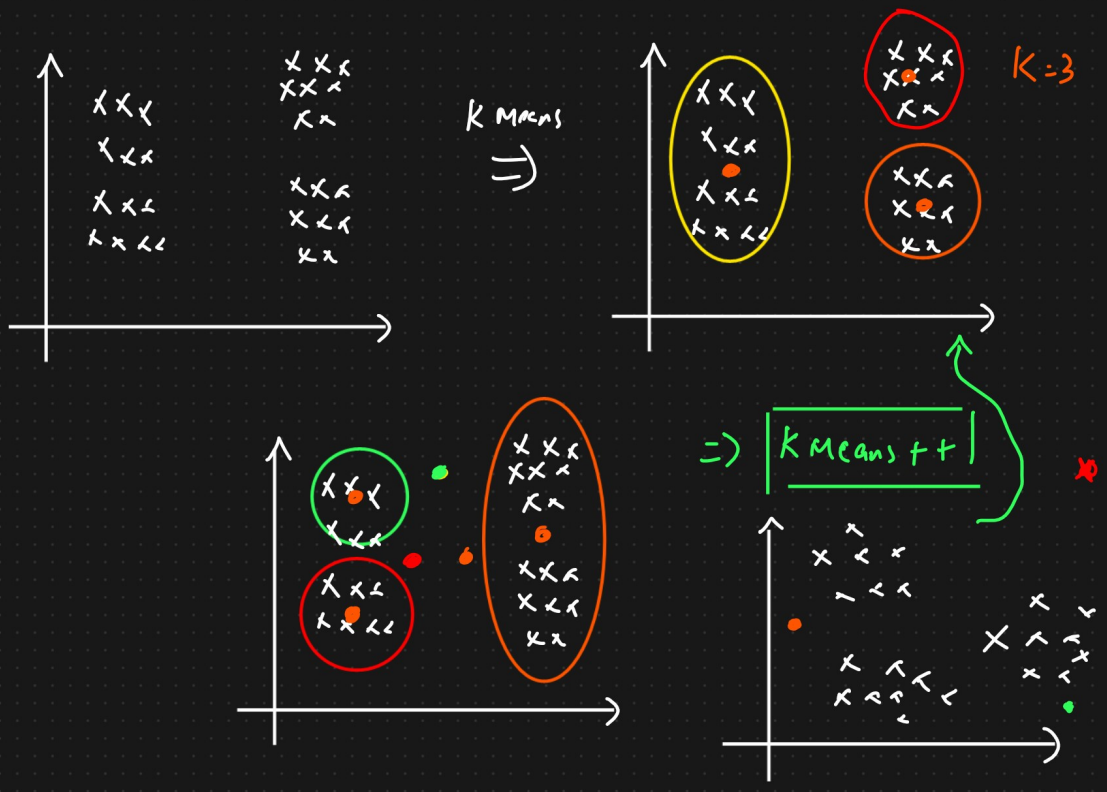
Euclidian Distance or Manhattan Distance



$$ED = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

$$MD = |x_2 - x_1| + |y_2 - y_1|$$

Random Initialization Trap (K Means ++)



K Means ++ Initialization Technique