

Full Name: Sajal Shrestha

CS622 ML Project-2

1. Using the following training data provided in nearest neighbors 1.csv, how would your algorithm classify the test points listed below with K=1, K=3, and K=5?

Tests	K=1	K=3	K=5
test1 =(1,1,1)	1	1	1
test2 =(2,1,-1)	-1	1	1
test3 =(0,10,1)	1	1	1
test4 =(10,10,-1)	-1	-1	-1
test5 =(5,5,1)	1	-1	1
test6 =(3,10,-1)	-1	1	-1
test7 =(9,4,1)	1	-1	-1
test8 =(6,2,-1)	-1	1	1
test9 =(2,2,1)	1	1	1
test10 =(8,7,-1)	-1	1	-1

2. What is the best K value for the training data above?

- When K=1.

3. Test your function on the following training data provided in clustering 2.csv, with K=2 and K=3. What changes do you notice when updating the k value?

- Whether K=2 or 3 in clustering-2.csv, the KMeans remains the same if the 'mu' array structure is not changed.

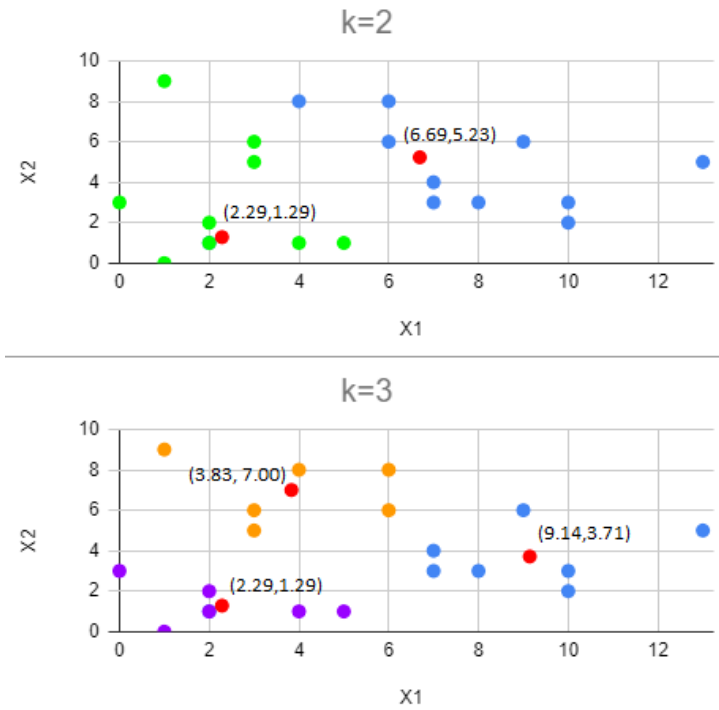
For example,

when K=2 or/and K=3 and mu =  $\begin{bmatrix} 1,2 \\ 5,2 \end{bmatrix}$  then, KMeans:  $\begin{bmatrix} 2.3 & 2.9 \\ 8. & 4.8 \end{bmatrix}$

when k=2 and mu =  $\begin{bmatrix} 1,2 \\ 5,2 \end{bmatrix}$  then, KMeans:  $\begin{bmatrix} 2.3 & 2.9 \\ 8. & 4.8 \end{bmatrix}$

when k=3 and mu =  $\begin{bmatrix} 1,2 \\ 5,2 \\ 2,3 \end{bmatrix}$  then, KMeans:  $\begin{bmatrix} 2.28571429 & 1.28571429 \\ 9.14285714 & 3.71428571 \\ 3.83333333 & 7. \end{bmatrix}$

4. Test your function with  $K=2$  and  $K=3$  on the above data. Plot your clusters in different colors and label the cluster centers.



5. Train your perceptron on the following dataset provided in perceptron 2.csv. Using the  $w$  and  $b$  you get, plot the decision boundary.

