Assignment 2

Ivan Efremov

November 2021

1 Task 1. Theoretical question on K-means Clustering

We can get 2 cases for cluster splitting:

1) 2m points belongs to first cluster and α belongs to second cluster.

2) Groups the m samples at x = 0 with the one at x = a (i.e., D2 = 0, ..., 0, a)

In the first case error $J_1 = m * (0-1)^2 + (-2-(-1))^2 + (a-a)^2 = 2m$
In the second case $J_2 = (-2-(-2))^2 + m*(0-0)^2 + (a-\frac{a}{m+1})^2 = (a-\frac{a}{m+1})^2$

So we want to get the second case, so the condition $J_2 < J_1$ must be satisfied.

$$J_2 < J_1 \Rightarrow (a - \frac{a}{m+1})^2 < 2m \Rightarrow a^2 < \frac{2*(m+1)^2}{m} \tag{1}$$

So

$$f(m) = \frac{2*(m+1)^2}{m} \tag{2}$$

2 Task 2. Theorectical question of SVM

I:

a) Yes

b) No. Because there is a shift, but in our case θ_0=0

c) No. We have a strict margin.

II:
a) Yes
b) Yes.
c) No. We have a strict margin.

III:
a) Yes
b) No. Because there is a shift, but in our case θ_0=0
c) Yes.