CS 6375

ASSIGNMENT 2 K-Means Clustering

Names of students in your group:

Naxel Santiago NAS180011

Danny Bao DXB180034

Number of free late days used: 0   
Note: You are allowed a **total** of 4 free late days for the **entire semester**. You can use at most 2 for each assignment. After that, there will be a penalty of 10% for each late day.

Please list clearly all the sources/references that you have used in this assignment.

1. **Prove that provided you make the following assumptions:**

Consider the formula for the error using the aggregated model:

We can move the outside of the brackets since it is a constant:

Using the Linearity of Expectation property, we can rewrite this as:

Again, using Linearity of Expectation, we can move the inside the summation:

Since (assumption 2), this can be simplified to:

This can be rewritten as:

Since , we can replace that with :

Thus, we have proven that .

1. **Show that using Jensen's inequality, it is still possible to prove that:**

Recall the formula . Since is a convex function, we can apply Jensen’s rule to the section of inside the brackets:

Take expectation of both sides:

Thus, we have proven that the inequality can still hold true even when the errors are correlated.

1. **Prove that at the end of *T* steps, the overall training error will be bounded by:**

Recall the form. This can be rewritten as:

Since :

Since is a distribution, if we sum both sides from *1* to *N*:

Now, consider the following formula for the training error of hypothesis :

Given , we can rewrite this as:

Recall the equation previously calculated for , and that if and if , where in this case. Thus, we can say that:

Which is equivalent to:

Now solve for :

Since , we can rewrite this as:

Recall the following formula for the training error of (from the lecture slides):

Using this, we can rewrite as:

Since :

Since for the total error of :

Using the inequality , we can conclude that:

Earlier we showed that . Thus, we now have:

Using the transitive law (which states that ), we can conclude that:

Thus, we have proven that the overall training error of the hypothesis will be less than or equal to .