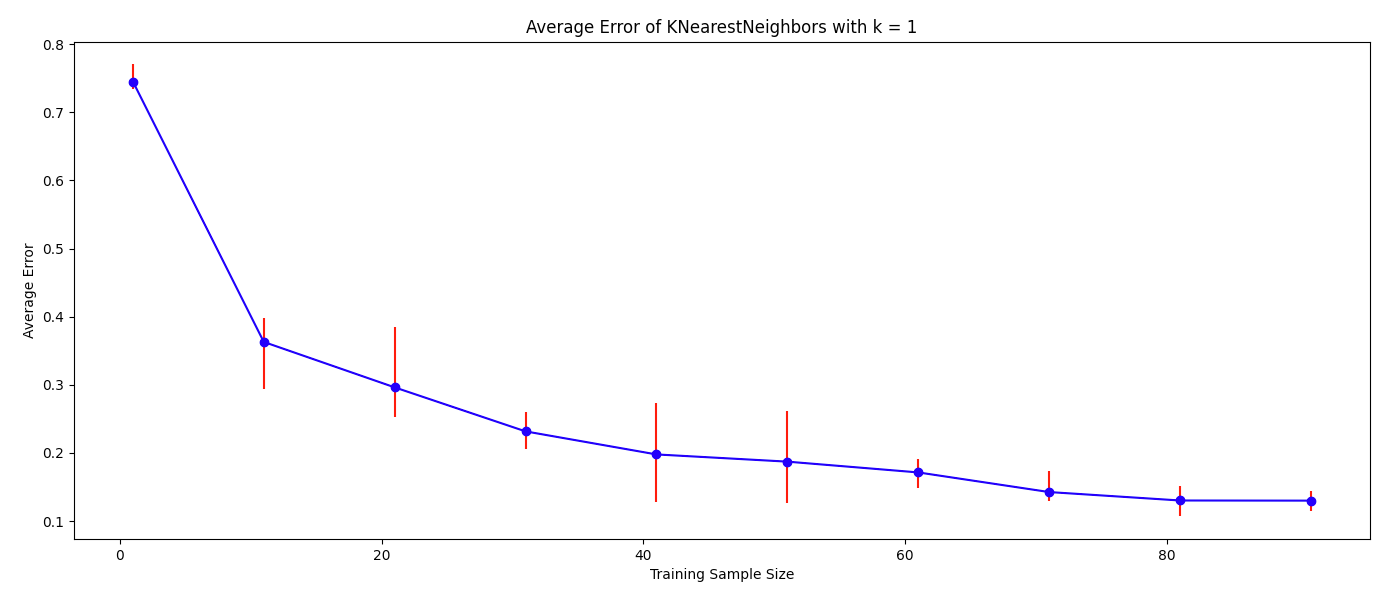
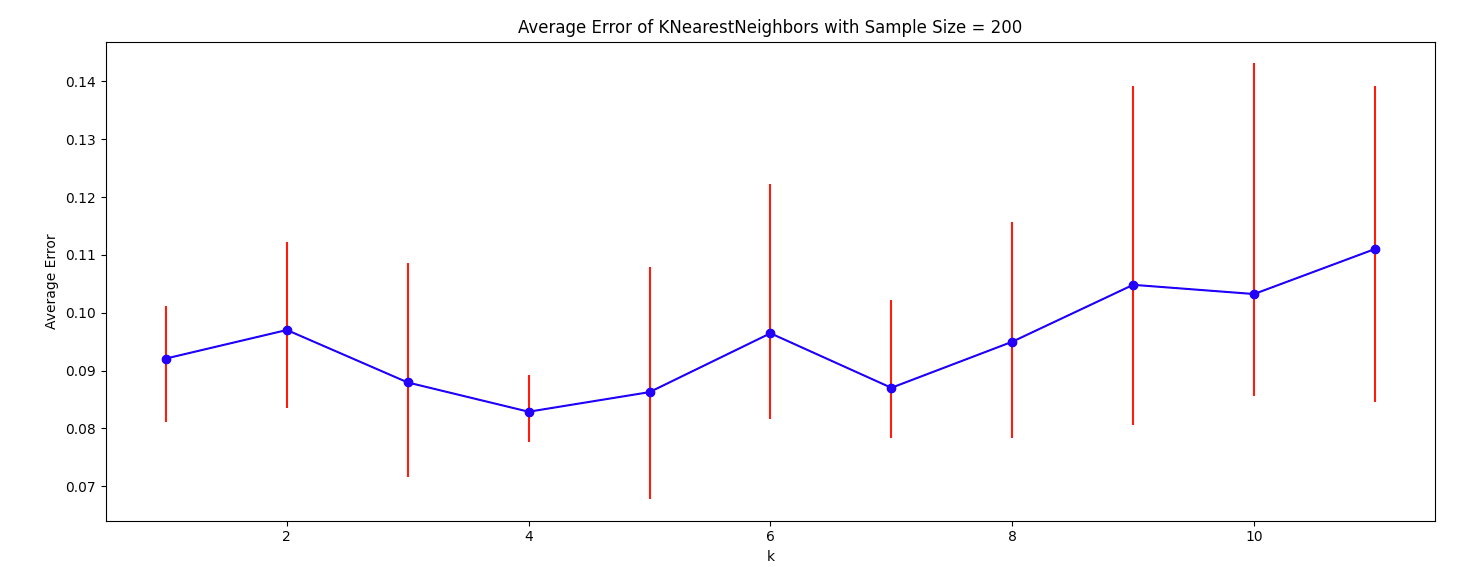
Exercise 1

Omri Bar Oz – 313325961, Ilana Pervoi – 318271640

Question 2:



1. We can observe that when the training sample size increases, the average test error decreases. It happens because when we increase the sample size, our samples become a better representation of the distribution and each point in the test set has a higher probability to find a closer neighbor from the samples with the same label during the NN algorithm run.
2. Yes. We get different results in different runs with the same sample size, since in each run the samples are chosen randomly, meaning they are not the same as in previous runs and therefore we might get a sample set with only 1 or 2 labels, which will cause us more errors than samples that are diverse and more representative of our distribution.
3. Yes, generally the error bars tend to decrease in size when the sample size increases. However, as we can observe there are cases in which the sample size increases, and yet the error bars increase, because as stated in section c, the accuracy of the model depends on how well the samples represent the distribution.



1. TODO
2. TODO

Question 3:

1. Need to prove:  
   for any two pairs if then

Proof:  
 and so WLOG we can decide that and .  
We obtain:

Recall that and of is c-Lipschitz with respect to the Euclidean distance, hence from a property we learned in class:

1. Need to prove:  
   under the given assumptions,

Proof:

So, we will show that .

Let and such that .  
Since is in at least one ball in the set of balls of radius that covers the space of points and there is an assumption that has a point in this ball as well, we obtain

Now, from the assumptions on we know that and has the same label

Question 4: