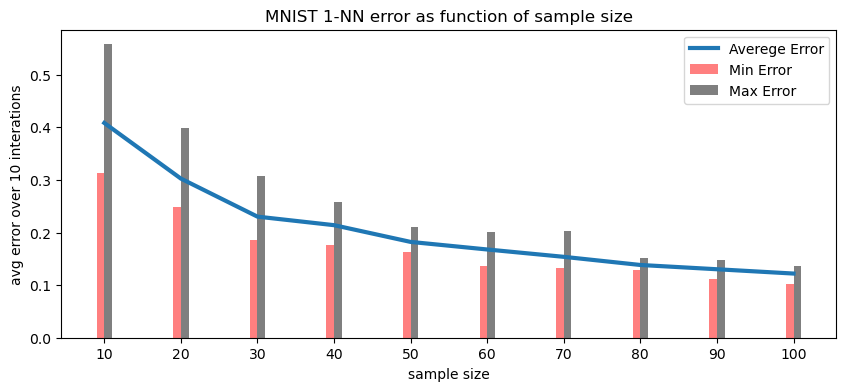
**Question 2:**

**a.**



**b.**

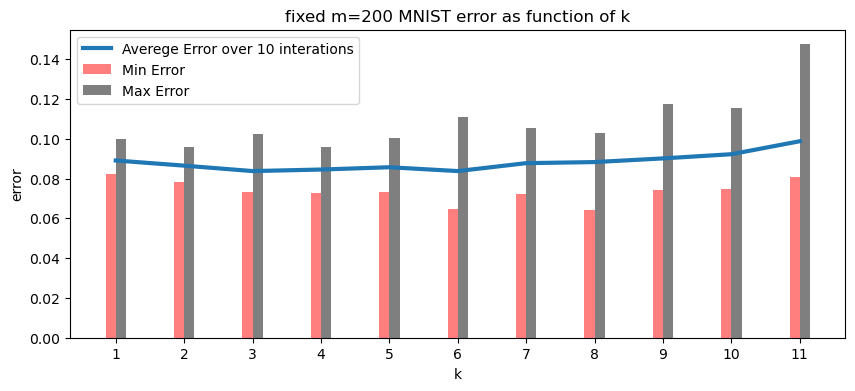
yes. The average error over 10 random sample decreases as the sample size increase This trend stems from the fact the when supplied with more examples and data, NN algorithm can make more precise generalization for the distribution based on the sample and thus generate a better rule.

**c.**

yes. we can see the difference between maximum and minimum error for each fixed sample size. this is because the algorithm learns and then tested each time on different samples so the prediction rule it generates changes.

**d.**

yes. We can see the gap between the maximum and minimum error shrinks as the sample size grows. When the algorithm sees more examples it can generate a better generalization on the distribution and even for special cases makes smaller errors.

**e.** 

**Question 3:**

a.

proof:

η of D is c-Lipschitz with respect to the Euclidean distance.

therefore, by definition: .

Hence, it’s sufficing to show that This is clear since

b.

Suppose for some there’s such as and

.

)

We know that covers the space of points in in balls size such that for every ball there is some pair which satisfies .

so

We also know that D is c-Lipschitz with respect to the Euclidean distance.

So and oppose to c-Lipschitz that

The support of D does not include two points with different  
labels that are less than -far.

Question 4:

a.

we can represent each rabbit as vector in since we only consider the parameters age and weight. it’s also known rabbits are limited to live 48 month and weigh 4kg, hence .

we would like to predict if rabbit is black or white so we can let or where means black rabbit.

b.

sd

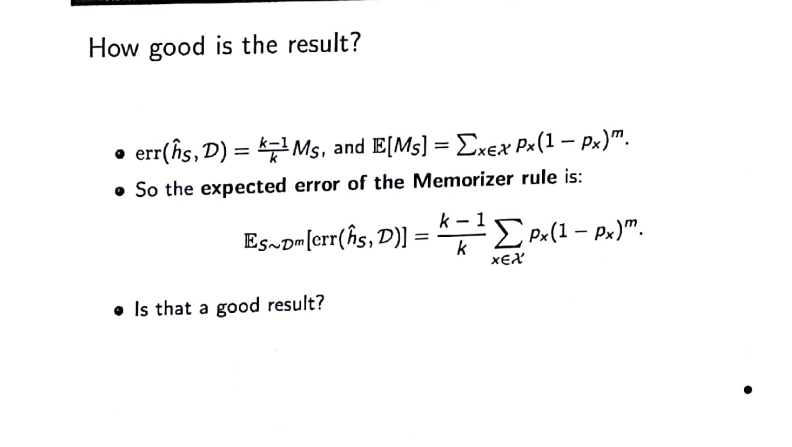
c.

d.

e.

f.

g.

[](file:///G:\My%2520Drive\uni\Machine%2520Learning%2520intro\סיכום%2520אלבוחר.pdf#page=6)