Thursday 10.11.2022

looked into different ways to make the data comparisons since I'm not sure if Im not sure if im allowed to get the euclidean distance between data points. (looked into hamming distance and also fourier transforms). I decided to wait until I get clarification before starting with that.

Friday 11.11.2022

- Struggled with the minks data for a very long time since I was going to use the tensorflow library to just get started but for some reason it did not want to install. After a long fight I threw in the towel and decided to go with csv files. I also looked into K-d trees

Saturday 12.11.2022

Set up the testing framework for the project and created a class that converts the mnist data files into csv files. Had some issues with reading and processing the csv files due to python having a limiter on memory usage for objects. So after trying to resolve that issue for a long time I decided to just decrease the training data so at the cost of accuracy. Sadly I had to cut the day short for personal reasons.

This week even though I encountered a few humps here and there, I feel like Im still on track with the project. Next i'll fully implement the unit testing and the Knn -algorithm (which oddly enough seems pretty straight forward, optimization seems to be the part where things are gonna get more difficult).

Also I was a bit confused with las weeks feedback:

- Huomaa, että luokittelussa tulee käyttää sen kaltaisia etäisyysmittoja kuin aihe-ehdotuksissa esitetään. Jos esim. laskee linkittämäsi artikkelin mukaisesti euklidisen etäisyyden harmaasävykuvien välillä, niin sillä saa jonkinlaisia tunnistustuloksia aikaan, mutta sellaisia luokitteluja on tehty kurssien yksittäisinä laskaritehtävinä.

from last week as I was not sure if i'm supposed to write down what method i will use for determining how close a point is to a group. And it was a bit unclear to me if I'm allowed to use euclidean distance or not since it has been present in prior tira tasks.