

Project Specification

TKT20010

Overview:

The project's goal is a program that will be able to recognise handwritten digits.

(If time allows then it would be nice with an option for the user to paint numbers on a canvas that then the program could identify)

Language:

- The project will be written in python. The languages that I can peer review are the following: C++, c# and python
- All code, weekly reports and Documentation, will be written in English.

Algorithms and Data structures:

- The algorithm this project will be built upon would be KNN (K nearest neighbor), the reason for choosing this algorithm is because it seems like a good entry point into machine learning.
- Using the Modified Hausdorff distance between black the pixels

Time/space complexity goals:

- Training time complexity: $O(d * n * \log(n))$
- Training space complexity: $O(d * n)$
- Prediction time complexity: $O(k * \log(n))$
- Prediction space complexity: $O(1)$
- Input:
- Data From the MNIST handwritten digit database of handwritten digits. Will be used both for training and testing the algorithms ability to recognise digits.

Degree programme:

- TKT

Sources:

<https://medium.com/analytics-vidhya/a-beginners-guide-to-knn-and-mnist-handwritten-digits-recognition-using-knn-from-scratch-df6fb982748a>

https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm

<https://towardsdatascience.com/k-nearest-neighbors-computational-complexity-502d2c440d5>

<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=576361>