Report for exercise 4 from group H

Tasks addressed:	4
Authors:	Ahmad Bin Qasim (03693345)
	Kaan Atukalp (03709123)
	Martin Meinel (03710370)
Last compiled:	2019–12–02
Source code:	https://gitlab.lrz.de/ga53rog/praktikum-ml-crowd

The work on tasks was divided in the following way:

Ahmad Bin Qasim (03693345)	Task 1	33%
	Task 2	33%
	Task 3	33%
	Task 4	33%
	Task 5	33%
Kaan Atukalp (03709123)	Task 1	33%
	Task 2	33%
	Task 3	33%
	Task 4	33%
	Task 5	33%
Martin Meinel (03710370)	Task 1	33%
	Task 2	33%
	Task 3	33%
	Task 4	33%
	Task 5	33%

Report on task 1, Principal component analysis

First part:

Energy of the two components: energy of pc1: 0.9929871280253524 energy of pc2 0.0070128719746476165

Second Part: Rerun it using columns rather than rows

Third Part: They are walking in loops. energy of pc1: 0.47330561274983274 energy of pc2 0.3759408098565421 Total Energy is around 84%. Two components are enough to capture the most energy. It is enough because in comparison to the original plot the reconstructed plot looks similar and still captures the most information.

Report on task 2, Diffusion Maps

Part 2:

Energy values: pc3 0.28867937283177014 pc2 0.3292051971173148 pc1 0.3821154300509151

The first two Principal components cover only 71.1% So we would lose 28.9% of the data. In comparison to that we can cover almost the whole dataset by taking the first three principal components. Consequently, it makes much more sense to take all principal components instead of two.

Report on task 3, Training a Variational Autoencoder on MNIST

What activation functions should be used for the mean and standard deviation of the approximate posterior and the likelihood and why?

Relu limits to positive values Tanh limits to positive values as well Sigmoid function function output is limited from -1 to 1 and maps all input to the borders. The latent space would be limited. This is why we use linear activations.

What might be the reason if we obtain good reconstructed and bad generated pictures? One reason if vanilla auto-encoders are used instead of variational auto-encoders. If the KL divergence loss is too high, then p_theta is not mapped to the Gaussian posterior distribution.

Report on task 4, Fire Evacuation Planning for the MI Building