

Exercise sheet 11

SoSe2025

Prof. Dr. Holger Fröhlich, Mohamed Aborageh, Anna Weller, Jiajun Qiu

Due date: July 1st

Exercise 1 - Variational Autoencoders (VAEs) (13 points)

The objective of this exercise is to train both a simple Autoencoder (AE) and Variational Autoencoder (VAE) on the Fashion-MNIST dataset and to compare their latent representations and reconstruction behaviors.

1. Load the Fashion-MNIST dataset from PyTorch datasets API. Binarize your input images, then flatten them into vectors of size 784. Use the following model architecture for the autoencoders:
 - Encoder: 2 hidden layers with ReLU and 2-dimensional latent space
 - Decoder: 2 hidden layers with ReLU, output layer with sigmoid activationCompile your models using Adam optimizer with a learning rate of 0.01. For AE, use binary cross entropy loss function and for VAE use ELBO with binary cross entropy as the likelihood term. Train your models using 20 epochs and a batch size of 128. Apply your models on 10 images from the Fashion-MNIST dataset, then plot the original and reconstructed images for reference. **(8 points)**
2. For a given image, encode and decode it multiple times with the AE and the VAE. What do you observe about the consistency of the outputs from each model? Justify your observations by considering and explaining the differences in the latent space representations used by a traditional AE and a VAE. **(3 points)**
3. Think of a way to visualize the data in the latent space. What do you observe when you distinguish the data points according to their labels? **(2 points)**

Exercise 2 – Optuna hyperparameter optimization (4 points)

Optuna is an open source hyperparameter optimization framework to automate hyperparameter search.

1. Use Optuna's Tree-structured Parzen Estimator (TPE) sampler to optimize the number of training epochs, batch size and learning rate for your VAE model from exercise 1 with 3-fold cross validation. Optimize for the lowest average binary cross entropy loss across the 3 folds using the following search spaces:
 - a. Epochs [10, 20, 30]
 - b. Batch size [64, 128, 256]
 - c. Learning rate [0.01, 0.1]Run 5 trials and report the best hyperparameter set and corresponding cross-validated loss. **(3 points)**
2. What is the difference between TPE, random search and grid search? **(1 point)**

Exercise 3 – Generative Models (8 points)

1. How can a VAE be used to generate new images? Demonstrate this by creating 20 new images using your trained VAE from exercise 1. **(2 points)**
2. What are the key differences between VAEs and GANs in terms of their loss functions and training objectives? What are the strengths of each approach? **(2 points)**
3. Research several notable GAN architectures that improve upon or differ from the original vanilla GAN. Select at least two different architectures and briefly explain how their techniques differ from the vanilla GAN. Provide a concrete example from the literature where this type of GAN has been used and cite the corresponding source **(4 points)**