Report of HW2 (Unsupervised Deep Learning)

Alireza Molla Ali Hosseini
(Dated: July 2022)

INTRODUCTION

Tasks

- Implement and test Auto-Encoder
- Optimize Hyper-parameters
- Fine-tune the Auto-Encoder by using a supervised classification task(compare it with the previous homework)
- Explore the latent space structure
- Implement and test Variational Auto-Encoder

The Goals of this homework is written as tasks above and the basic strategies is to define an Auto-Encoder. In order to do so we should have an Encoder part and a Decoder part which both of them are convolutional neural networks. The Encoder part tries to encode the data into a latent space with a specified dimension. Then the Decoder part is trying to decode the latent space to actual space of data and reproduce them. The result of Auto-Encoder network is in Table V

In this homework I used the Optuna to optimize the hyper-parameters. And after finding the best ones and creating an Auto-Encoder with those hyper-parameters, the results were driven by Cross-Validation method (which divide the data set into k (k=10) parts and use k-1 of them as training data and one of them as validation and this procedure happens k time. Then the result will be the average of all the k results). Table V

In Auto-Encoder method we won't need the label of the data(it is an unsupervised method), but since the data are Fashion MNIST Dataset the labels are also available. Therefor we can use them to appreciate the accuracy of the unsupervised method (V). In order to do so, we should encode the data with their related labels by Encoder and then try to classify the latent space into 10 classes with TSNE method(since we have the labels). Then if we take a random sample from those classes and decode them into real space, we can plot the picture and see the result(the anticipated class). Figure 3

METHODS

Encoder

It is a convolutional neural network which contains three convolutional network followed by two fully connected network. The inputs are two-dimension pictures of Fashion MNIST (as tensor type) and the output is a tensor of specified latent space dimension. The parameters of the network are summarized in Table I and the latent space is plotted in Figure 3.

Decoder

It is also another convolutional network which contains 2 layers of fully connected network followed by three transposed convolution (reverse of convolution). The inputs are tensor of latent space and the output is two dimension pictures(reproduced pictures). The parameters of the network is summarized in Table II and the latent space is plotted in Figure 3.

Optuna

Optuna is used to find the best hyper-parameters of this task. To do so, first we need to define a model that optuna will have for different trials. Then we should specify different hayper-parameters that we want them to be tuned. Then after a fixed number of trials we will have the best trial with the best hyper-parameters. In this task the hyper-parameters are latent space, opimizing method, learning rate and regulation rate. The Best hyper-parameters of Optuna are summarized in Table VI and the related latent space of the network with this hyper-parameters is plotted in Figure 3.

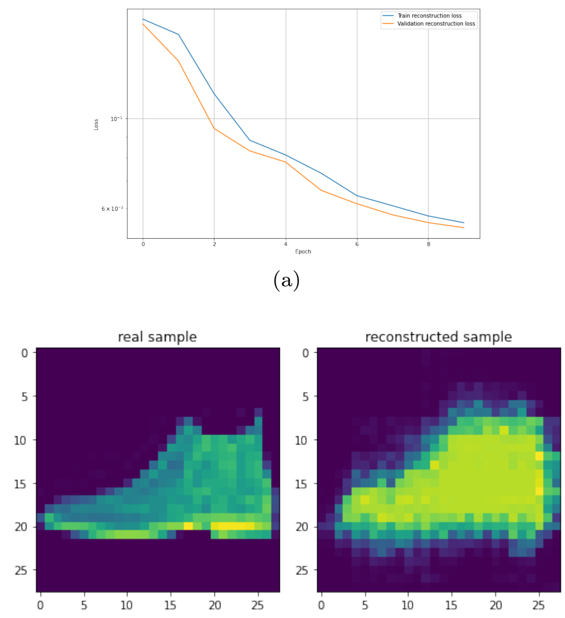
Fine-Tuning

In this method we can fine-tune an unsupervised network with a supervised network to measure the performance of the unsupervised network. This method, obviously, available only if we would have the labels of our Dataset. The parameters of this network is summarized in Table III and the latent space in plotted in Figure 3.

Variational-AutoEncoder(VAE)

Variational Auto-Encoder is another architecture in neural network which has similarity with Auto-Encoder. However, in a VAE the input data is sampled from a parametrized distribution (z) and the encoder and decoder are trained jointly such that the output minimizes a reconstruction error in the sense of the Kullback–Leibler divergence between the parametric posterior and the true posterior. The parameters of this network is summarized in Table IV and the latent space is plotted in Figure 3

RESULTS



(a)

(b) (c)

FIG. 1: (a) Auto-Encoder Loss (b) Auto-Encoder Input sample (c) Auto-Encoder Reconstructed sample

(a)

(b)

FIG. 2: (a) VAE sample Input Batch Visualization (b) VAE sample Reconstructed (20 Epoch)

Parameter	Value
Batch size	256
Number of Epochs	10
Latent Space Dimension	2
Convolutional Layers	3
Linear layers	2
Activation function	**ReLU**
Optimizer	**Adam**
Loss Function	**MSE**
learning rate	5e-4
weight decay	1e-5
Kernel	3
Stride	2
Padding	1

TABLE I: Encoder Parameters

Parameter	Value
Batch size	256
Number of Epochs	10
Latent Space Dimension	2
Transpose Convolutional Layers	3
Linear layers	2
Activation function	**ReLU**
Optimizer	**Adam**
Loss Function	**MSE**
learning rate	5e-4
weight decay	1e-5
Kernel	3
Stride	2
Padding	1

TABLE II: Decoder Parameters

Parameter	Value
Batch size	256
Number of Epochs	20
Number of Folds (Cross-Validation)	10
Linear layers	3
Activation function	**ReLU**
Optimizer	**Adam**
Loss Function	**Negative Log Likelihood**
learning rate	0.02
weight decay	1.5e-5
Drop out	0.1

TABLE III: Fine-Tune Parameters

Parameter	Value
Batch size	256
Hidden Space Dimension	64
Latent Space Dimension	10
Convolutional Layers	3
Linear layers	2
Activation function	**ReLU**
Optimizer	**Adam**
Weight initialization	**Kaiming**
Sample Distribution	**Normal Distribution**
learning rate	1e-3
weight decay	1e-5
Number of Epochs	20
Kernel	3

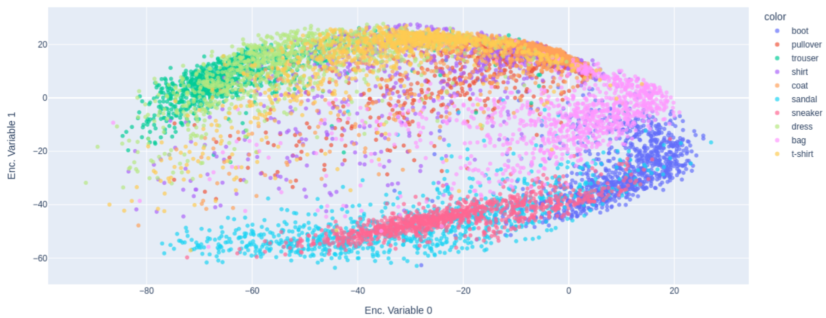
TABLE IV: VAE Parameters

Network	Accuracy
Auto-Encoder	94.64
Auto-Encoder (Optuna)	98.27
Fine-Tune	9.7
VAE(after 20 epoch)	243(Loss)

TABLE V: Networks Results

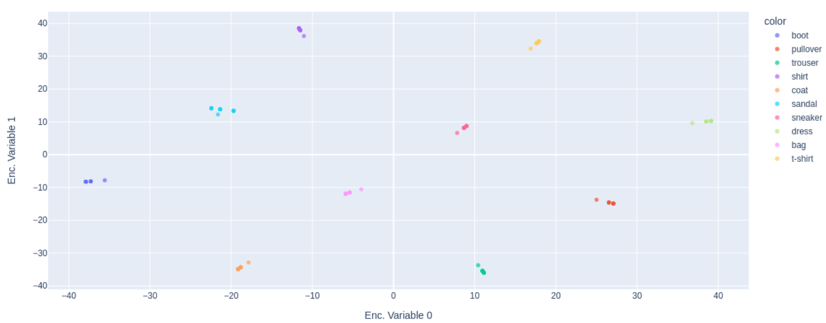
Parameter	Range	Best
Latent Space Dimension	2-20	11
Learning rate	1e-5 - 1e-1	0.02
Weight Decay	1e-5 - 1e-1	1.54e-5

TABLE VI: Optuna search results for Auto-Encoder network

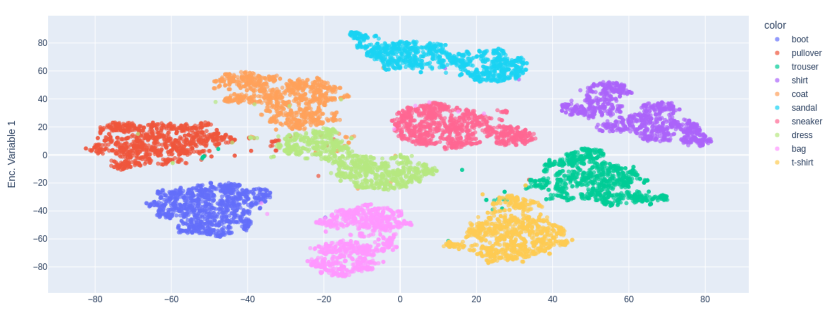


(a)

(b)



(c)



(d)

FIG. 3: Latent space of: (a) Auto-Encoder (b) Auto-Encoder(Optuna) (c) Fine-Tune (d) VAE