

CSC7343 HW1

Due Oct 8th by the end of the day. Do all the work (code, plot, analysis text) on a jupyter notebook. On the notebook, you should clearly mark the code segment that implements the models, i.e., CNN, VAE and the combined. Upload the notebook with all the results/text in moodle on or before the due time. [All implementation should use PyTorch.]

(Also upload the weights of the trained model to your google drive and put the share link on the notebook. Make sure I can download the weights and load into the models on the notebook to repeat your results. Put all your code for training the model in one cell so that I can disable them and load weights from you file directly.)

Your tasks in this homework are to explore CNN classification and autoencoders.

Task 1: Implement a CNN with the following structure.

Softmax(10)
FC(256)
Conv(128) filter size: 3x3, stride: 2
Conv(64) filter size: 3x3, stride: 2
Conv(32) filter size: 3x3, stride: 2

Train and test the network on fashion mnist dataset (train using train data and test using test data in the dataset). Use cross-entropy loss and an optimizer of your choice. Train the network until the loss converges.

1. Collect train and test accuracy about every 10 epochs during the process. Plot the two accuracy numbers against the epochs.

Task 2: Implement a variational autoencoder (VAE) whose encoding part should have the same structure as the above CNN (i.e. 3 conv layers and an encoding of length 256). The decoding part should have a mirror structure of the encoder (i.e., the input/output shapes of a decoding layer should be the inverse of the shapes at the corresponding encoding layer. Use ConvTranspose2D as reverse of Conv2d). You may use squared error or binary cross-entropy as reconstruction loss. Train the VAE with the fashion mnist training data until convergence.

1. Plot the distribution of the 2-norms of the encoding vectors
2. Plot 3-4 original images and the reconstructions side by side to exam how good the reconstructions are.

Task 3: Investigate whether combining VAE with the CNN can improve the CNN's classification. Modify the VAE such that the encoding (the mean) is passed to a softmax layer for classification. Train the joint model with the combined loss: cross-entropy from the supervised component together with reconstruction and Kullback–Leibler divergence loss from the AE component. Train the model with fashion mnist data until convergence.

1. Collect train and test accuracy about every 10 epochs during the process. Plot the two accuracy numbers against the epochs
2. Is the test accuracy of the joint model better than the CNN in task 1? If so, why? If not, why not?
3. Alternatively, one can train a VAE first and then replace the decoder with a softmax layer to form a classifier. Afterwards, train the classifier with the labeled data. Try this approach and compare it with the joint model. Which one gives better performance on testing data?