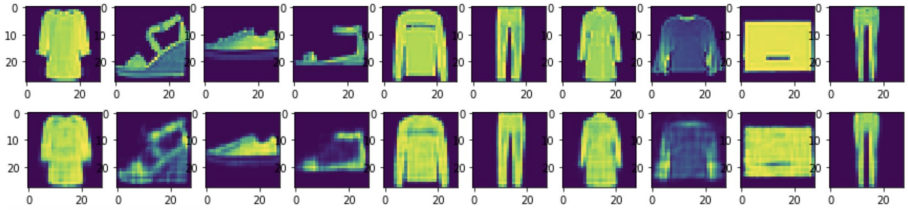
CSCI-6960 ML & Optimization HW6

Dong Hu

1



numepochs = 20, numsplits = 400, $\alpha = 0.001$, $\rho_1 = 0.9$, $\rho_2 = 0.999$, $\delta = 10^{-8}$

Training loss VS epoch:
Training loss [1/20]: 0.70
Training loss [2/20]: 0.30
Training loss [3/20]: 0.28
⋮
Training loss [17/20]: 0.26
Training loss [18/20]: 0.26
Training loss [19/20]: 0.26
Training loss [20/20]: 0.26

Final loss for testing data: 0.26661104871109065

Observations: For the training part, the loss converges very fast and remained around 0.26 for different combinations of (numepochs,numsplits), if we allow more epochs for training it gives slightly less loss for training data but doesn't really help improving the loss for testing. the plot generated by the autoencoder seems to be blurred a little bit compared to the original image, so I think it's doing something similar to what PCA/Low rank Approximation doing to keep the most important infomation and get rid of the redundant one while training.