HWK5

1. Let consider a training set that consists of 4 labeled data points:

$$\left(\begin{bmatrix} -1 \\ 2 \end{bmatrix}, +1\right) \quad, \quad \left(\begin{bmatrix} -1 \\ 5 \end{bmatrix}, +1\right) \quad, \quad \left(\begin{bmatrix} -1 \\ -1 \end{bmatrix}, -1\right) \quad, \quad \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, -1\right)$$

So for example the third data point is the vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and it belongs to category -1. Consider a two layers network, without bias, with a ReLU between the two linear layers, and with a Hinge Loss Criterion:

$$\text{Linear} \to \text{ReLU} \to \text{Linear} \to \text{Hinge Loss}$$

The weights for the first and second layer are:

$$U = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} \frac{1}{4} & -\frac{1}{2} \end{bmatrix}$$

(a) Compute with pen and paper the average loss on the minibatch that consists of the first and last data point, as well as the average loss on the full training set. Show your computaion.

.

$$\mathcal{L}^{(1,4)} = \qquad\qquad \mathcal{L} =$$

(b) Follow the steps in the Jupyter Notebook "problem1.ipynb": you will implement the above neural net in Pytorch. Then you will store the training set into a tensor train_data and a tensor train_label. Then you will use the network to compute the average loss on the batch (1,4) and on the full training set. Obviously you should get back the same result than in part a).

2. Make a multiple layer neural net and try to get the best possible result on FASHION-MNIST. Try various number of layers, and various sizes for the hidden layers. Report the error rate on the train set as well as on the test set for your best neural net. Also report the number of parameters and training time. The student with the lowest error rate on the test set will get 3 bonus points on the take home exam.

   - Number of layers and hidden sizes:
   - Number of parameters:
   - Time it took to train the network:
   - Error rate on the training set:
   - Error rate on the test set:

3. Create a virtual machine with a Tesla K80 GPU (all the instructions are on brighspace).