1. Group Members:
   1. ag3743, Aman Gullati, section 001
   2. hs2923, Henrique Saboya Lopes Tavares de Melo, section 002
   3. js4637, Joaquim Lyrio, section 002
   4. rap2194, Ricardo Pommer, section 002
2. As we know, neural networks are considered to be a black box model by many people. What we want is to be able to understand what our neural network is doing at each layer and which information is it capturing. Thus, our problems will be to implement a feature visualization algorithm, interpret what feature map our neural network uses and how each layer reacts to the inputs.
3. We will be looking mainly at two datasets. First, we will look at a relatively simpler dataset of MNIST and understand how our neural network reacts, then we will look at a slightly more complex dataset of CIFAR 10 and see how feature visualization works in that case. The MNIST data set will have image inputs with the same dimensions, and 10 labels (0 through 9) as output, so that we can start with “baby steps”. The CIFAR 10 dataset will have more complex coloured images (tough with the same dimensions) as input, and also 10 labels as outputs, however, this time, the labels will be animals and objects.
4. Our starting point is something that we worked on in class. The feature visualization for classification of E, F & L. We will visualize the weight values of each layer in the MNIST dataset as if they were images, for a very simple neural network. This is feasible for smaller neural networks but too complex for larger neural networks. Then, our next step would be to look at the feature visualization as an optimization problem, understanding which image activates the network the most. After getting these 2 steps, we’ll be working on the CIFAR10 data, using more complex neural network structures. There are multiple methods of looking at feature visualization, at different difficulty levels, some of them are: using a deconvoluted neural network and using a regularized optimization in image space. We have been able to find some literature with detailed procedures for performing these methods.
5. The most difficult aspects of the project will be creating the feature visualization algorithm from scratch, handling the optimization of the images and the complexity of the model which may make running times quite high. We are looking at multiple papers which provide us optimization methodologies and we plan on running the code in a GPU to ensure that it runs at a good speed.