* ***Model & Training Procedure Description***
  + Experiment X
    - Number of Layers:
    - Number of Neurons:
      * Layer X:
      * Layer Y:
    - Number of Epochs
    - Mini batch size
    - Weight initialization scheme:
    - Activation function:
    - Accuracies obtained:
* ***Model Performance & Confusion Matrix***
  + Number of Layers:
  + Number of Neurons:
    - Layer X:
    - Layer Y:
  + Number of Epochs
  + Mini batch size
  + Weight initialization scheme:
  + Activation function:
  + Accuracies obtained:

Include a section describing in more detail the most accurate model you were able to obtain: the architecture of your model, including number of layers, number of neurons in each layer, weight initialization scheme, activation function, number of epochs used for training, and batch size used for training.

Include a confusion matrix showing results of testing the model on the test set. The matrix should be a 10-by-10 grid showing which categories images were classified as. Use your confusion matrix to additionally report precision & recall for each of the 10 classes, as well as overall accuracy of your model.

* ***Training Performance Plot***

For your best performing ANN, include a plot showing how training accuracy and validation accuracy change over time during training. Graph number of training epochs (x-axis) versus training set and validation set accuracy (y-axis). Hence, your plot should contain two curves.

* ***Visualization***

Include 3 visualizations of images that were misclassified by your best performing model and any observations about why you think these images were misclassified.  You will have to create or use a visualization program that takes a 28-by-28 matrix input and translate it into a black-and-white image.