* ***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.

Constant configuration:

* Weight Initialization Scheme: he\_normal
* Input Layer Activation Function: ReLu
* Output Layer Activation Function: Softmax
* Experiment 1
  + Number of Layers: 1 hidden layer(s)
  + Number of Neurons: 100
    - Layer 1: ReLu, 100 Neuron(s)
  + Number of Epochs: 1000
  + Mini batch size: 1000
  + Accuracies obtained: 11.19%
* Experiment 2
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 300
    - Layer 1: Sigmoid, 100 Neurons
    - Layer 2: Sigmoid, 100 Neurons
    - Layer 3: Sigmoid, 100 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 1000
  + Accuracies obtained: 77.66%
* Experiment 3
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 300
    - Layer 1: Sigmoid, 100 Neurons
    - Layer 2: Sigmoid, 100 Neurons
    - Layer 3: Sigmoid, 100 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 1000
  + Accuracies obtained: 38.86%
* Experiment 4
  + Number of Layers: 4 hidden layer(s)
  + Number of Neurons: 400
    - Layer 1: Sigmoid, 100 Neurons
    - Layer 2: Sigmoid, 100 Neurons
    - Layer 3: Sigmoid, 100 Neurons
    - Layer 4: Sigmoid, 100 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 1000
  + Accuracies obtained: 45.27%
* Experiment 5
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 300
    - Layer 1: Sigmoid, 100 Neurons
    - Layer 2: Sigmoid, 100 Neurons
    - Layer 3: Sigmoid, 100 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 512
  + Accuracies obtained: 80.78%
* Experiment 6
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 300
    - Layer 1: Sigmoid, 100 Neurons
    - Layer 2: Sigmoid, 100 Neurons
    - Layer 3: Sigmoid, 100 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 100
  + Accuracies obtained: 87.27%
* Experiment 7
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 300
    - Layer 1: Sigmoid, 100 Neurons
    - Layer 2: Sigmoid, 100 Neurons
    - Layer 3: Sigmoid, 100 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 87.57%
* Experiment 8
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 300
    - Layer 1: Tanh, 100 Neurons
    - Layer 2: Tanh, 100 Neurons
    - Layer 3: Tanh, 100 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 73.56%
* Experiment 9
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 300
    - Layer 1: Selu, 100 Neurons
    - Layer 2: Selu, 100 Neurons
    - Layer 3: Selu, 100 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 25
  + Accuracies obtained: 73.56%
* Experiment 10
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 300
    - Layer 1: Selu, 100 Neurons
    - Layer 2: Selu, 100 Neurons
    - Layer 3: Selu, 100 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 25
  + Accuracies obtained: 10.04%
* Experiment 11
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 300
    - Layer 1: Selu, 100 Neurons
    - Layer 2: Selu, 100 Neurons
    - Layer 3: Selu, 100 Neurons
  + Number of Epochs: 100
  + Mini batch size: 25
  + Accuracies obtained: 9.91%
* Experiment 12
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 300
    - Layer 1: Selu, 100 Neurons
    - Layer 2: Selu, 100 Neurons
    - Layer 3: Selu, 100 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 60
  + Accuracies obtained: 10.22%
* Experiment 13
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 300
    - Layer 1: Sigmoid, 100 Neurons
    - Layer 2: Sigmoid, 100 Neurons
    - Layer 3: Sigmoid, 100 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 87.03%
* Experiment 14
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 300
    - Layer 1: Selu, 100 Neurons
    - Layer 2: Selu, 100 Neurons
    - Layer 3: Selu, 100 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 250
  + Accuracies obtained: 9.18%
* Experiment 15
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 300
    - Layer 1: Sigmoid, 50 Neurons
    - Layer 2: Sigmoid, 200 Neurons
    - Layer 3: Sigmoid, 50 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 89.11%
* Experiment 16
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 150
    - Layer 1: Sigmoid, 50 Neurons
    - Layer 2: Sigmoid, 50 Neurons
    - Layer 3: Sigmoid, 50 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 88.07%
* Experiment 17
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 450
    - Layer 1: Sigmoid, 200 Neurons
    - Layer 2: Sigmoid, 50 Neurons
    - Layer 3: Sigmoid, 200 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 85.74%
* Experiment 18
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 450
    - Layer 1: Sigmoid, 200 Neurons
    - Layer 2: Sigmoid, 50 Neurons
    - Layer 3: Sigmoid, 200 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 85.74%
* Experiment 19
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 350
    - Layer 1: Sigmoid, 200 Neurons
    - Layer 2: Sigmoid, 100 Neurons
    - Layer 3: Sigmoid, 50 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 84.82%
* Experiment 20
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 350
    - Layer 1: Sigmoid, 50 Neurons
    - Layer 2: Sigmoid, 100 Neurons
    - Layer 3: Sigmoid, 200 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 88.49%
* Experiment 21
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 300
    - Layer 1: Sigmoid, 25 Neurons
    - Layer 2: Sigmoid, 250 Neurons
    - Layer 3: Sigmoid, 25 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 87.33%
* Experiment 22
  + Number of Layers: 3 hidden layer(s)
  + Number of Neurons: 200
    - Layer 1: Sigmoid, 50 Neurons
    - Layer 2: Sigmoid, 100 Neurons
    - Layer 3: Sigmoid, 50 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 87.70%
* Experiment 23
  + Number of Layers: 4 hidden layer(s)
  + Number of Neurons: 500
    - Layer 1: Sigmoid, 50 Neurons
    - Layer 2: Sigmoid, 200 Neurons
    - Layer 3: Sigmoid, 200 Neurons
    - Layer 4: Sigmoid, 50 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 84.94%
* Experiment 24
  + Number of Layers: 5 hidden layer(s)
  + Number of Neurons: 200
    - Layer 1: Sigmoid, 100 Neurons
    - Layer 2: Sigmoid, 200 Neurons
    - Layer 3: Sigmoid, 300 Neurons
    - Layer 4: Sigmoid, 200 Neurons
    - Layer 5: Sigmoid, 100 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 86.60%
* Experiment 25
  + Number of Layers: 5 hidden layer(s)
  + Number of Neurons: 450
    - Layer 1: Sigmoid, 50 Neurons
    - Layer 2: Sigmoid, 100 Neurons
    - Layer 3: Sigmoid, 150 Neurons
    - Layer 4: Sigmoid, 100 Neurons
    - Layer 5: Sigmoid, 50 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 86.11%
* Experiment 26
  + Number of Layers: 5 hidden layer(s)
  + Number of Neurons: 225
    - Layer 1: Sigmoid, 25 Neurons
    - Layer 2: Sigmoid, 50 Neurons
    - Layer 3: Sigmoid, 75 Neurons
    - Layer 4: Sigmoid, 50 Neurons
    - Layer 5: Sigmoid, 25 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 88.19%
* Experiment 27
  + Number of Layers: 5 hidden layer(s)
  + Number of Neurons: 120
    - Layer 1: Sigmoid, 10 Neurons
    - Layer 2: Sigmoid, 25 Neurons
    - Layer 3: Sigmoid, 50 Neurons
    - Layer 4: Sigmoid, 25 Neurons
    - Layer 5: Sigmoid, 10 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 59.42%
* Experiment 28
  + Number of Layers: 5 hidden layer(s)
  + Number of Neurons: 120
    - Layer 1: Sigmoid, 15 Neurons
    - Layer 2: Sigmoid, 25 Neurons
    - Layer 3: Sigmoid, 50 Neurons
    - Layer 4: Sigmoid, 25 Neurons
    - Layer 5: Sigmoid, 15 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 79.31%
* Experiment 29
  + Number of Layers: 5 hidden layer(s)
  + Number of Neurons: 220
    - Layer 1: Sigmoid, 15 Neurons
    - Layer 2: Sigmoid, 25 Neurons
    - Layer 3: Sigmoid, 150 Neurons
    - Layer 4: Sigmoid, 25 Neurons
    - Layer 5: Sigmoid, 15 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 81.82%
* Experiment 30
  + Number of Layers: 2 hidden layer(s)
  + Number of Neurons: 100
    - Layer 1: Sigmoid, 50 Neurons
    - Layer 2: Sigmoid, 50 Neurons
  + Number of Epochs: 1000
  + Mini batch size: 50
  + Accuracies obtained: 83.42%
* Experiment 31
  + Number of Layers: 5 hidden layer(s)
  + Number of Neurons: 220
    - Layer 1: Sigmoid, 15 Neurons
    - Layer 2: Sigmoid, 25 Neurons
    - Layer 3: Sigmoid, 150 Neurons
    - Layer 4: Sigmoid, 25 Neurons
    - Layer 5: Sigmoid, 15 Neurons
  + Number of Epochs: 10000
  + Mini batch size: 50
  + Accuracies obtained: 74.29%