

Q1.Explain your choice of neural network architecture: how many layers your network has? What types of layers does it contain? What about other decisions like use of dropout layers, activation functions, number of channels / hidden units.

The neural network has 1 input layer, 3 activation layers and 1 output layer. Dropout layers have not been used as the results were better without them. ReLU activation has been used everywhere to avoid vanishing gradient problem.

Q2. Explain your choice of loss function.

Cross entropy Loss function has been used as it is best suited to represent the overall distribution.

Q3. The weight decay and dropout layers were added to improve the validation accuracy

This was the best accuracy obtained by using the pretrained model

