SMAI NEURAL NETWORK ASSIGNEMENT

Manav Bhatia(2018102009)

Q1.

We created a Neural Network with 6 Layers since the input to the Network is quite large, we need a sufficient number of layers and learnable parameters to correctly identify the given gestures.

The first layer which is the input layer has (3*224*224) neurons, basically each pixel of the input image is a Neuron. The second layer has (128*3) Neurons, the third layer has 128 neurons, the fourth layer has 64 neurons, the fifth layer has 32 neurons and the output layer has 9 Neurons, basically the number of output classes.

We have 6 Linear Layers-

a. 1 input layer,
b. 4 hidden layers(with all the learnable parameters)
c. 1 output layer.

We have used ReLU Activation function with ease, since the inputs and outputs are both positive. Dropout layers with 40% dropout rate have been used later to counter overfitting.

Q2.

For the training step, we have used cross-entropy loss since it works best as multiclass log loss is minimised and rewards the best weights. Also, we use SGD since it generally provides us with the best results in Neural Network Architecture.

Q3.

A) We clearly observe the dataset overfitting to small training data as we achieve perfect training accuracy but very small validation accuracy of 0.35. Hence we observe very high overfitting

B)We have used various techniques like dataset augmentation, weight decay and Dropout to reduce the effect of overfitting.As we can clearly observe, the accuracy on augmented training set drops but then, it dies not perform very badly on Validation set either.

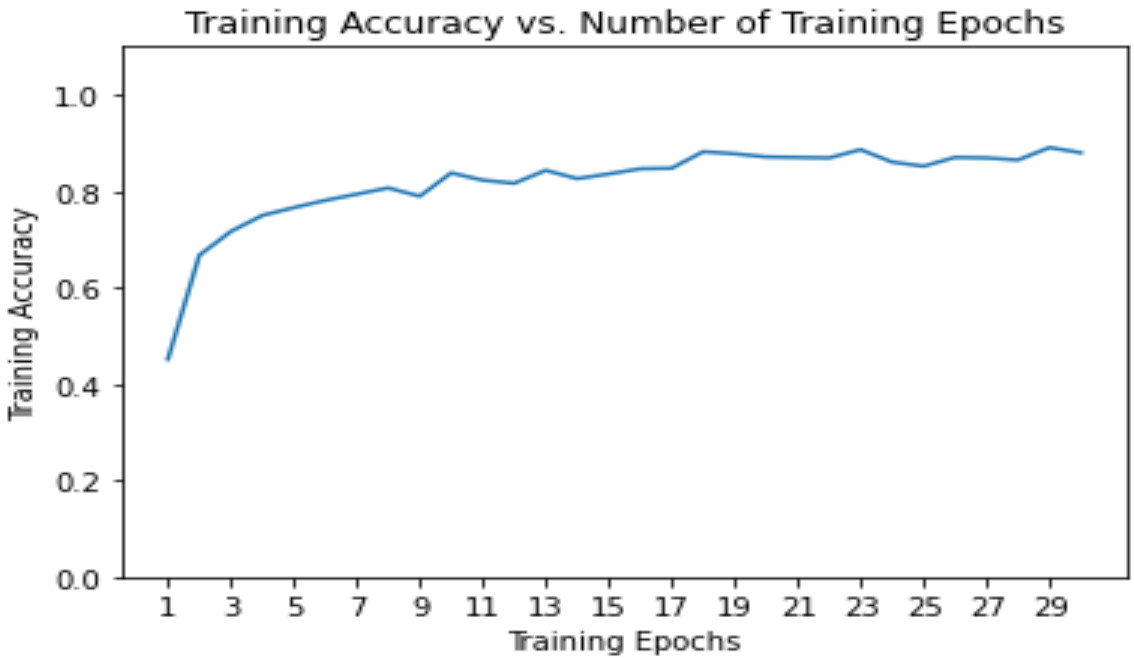
Q4.

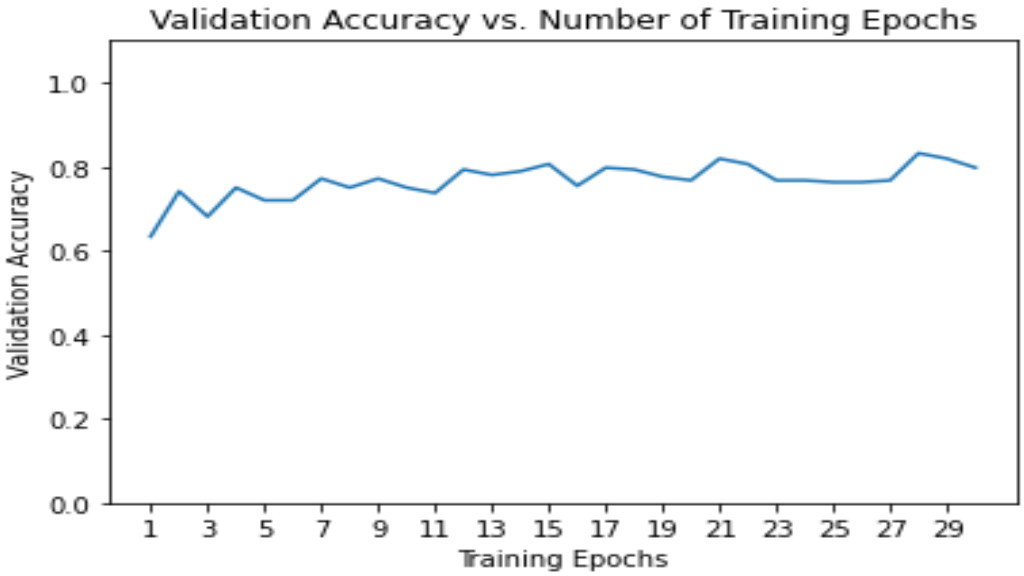
In this question, we have used a pre-trained AlexNet architecture and modified its last layer to get the same number of features as our output(9) requires and used cross-entropy and SGD as above. Alexnet is preferred as it is pre-trained on the Image-1000 dataset and is constructed very well to handle image based data.

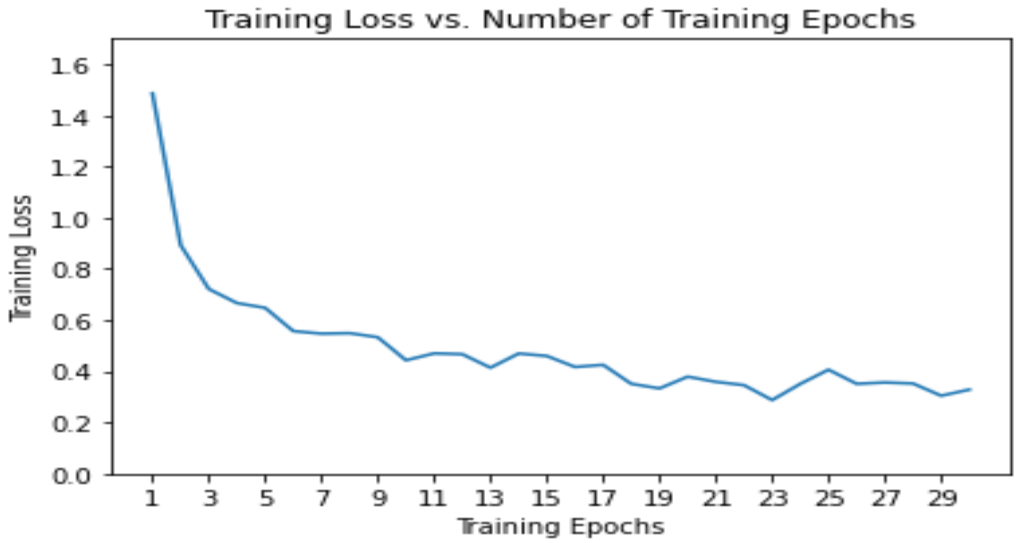
Q5.

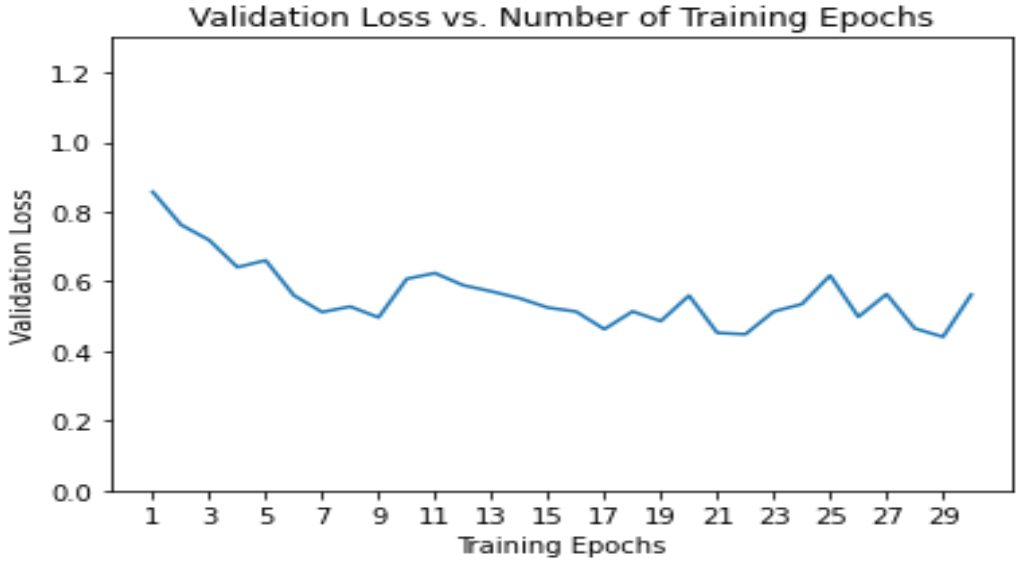
As was seen, the training accuracy and the validation accuracy drastically increase in the first few iterations but gradually settle down the longer we allow the model to train, that is, the first few steps show major weight changes but that gradually slows down and settle for high accuracy.

Similarly, we have high losses in the beginning stage as we require much changes but that gradually settles down the longer we train and test, bringing down the losses to a more stable version for both training and validation.


Training Accuracy vs. Number of Training Epochs
Training Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29
Training Epochs


Validation Accuracy vs. Number of Training Epochs
Validation Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29
Training Epochs





The final Validation Accuracy achieved by our best model is 83.1897%, a reasonable increase over the MLP accuracy of 17.2222% and a part of the reason why pre-trained models, a complex combination of Linear, Polynomial and Convolutional Neural Layers are preferred on Image-based Datasets.