Machine Learning Spring 2022
Assignment 3 Report

Ali Khalid
MSDS21001

May 27, 2022

1 Task # 1

1.1 Model Parameters

- Learning rate = 0.0001
- Optimizer = Adam
- Epochs = 20
- model Layer with neurons in each layer:
  - Input Layer: 3*218*178
  - Hidden Layers: 2X[Conv2d, BatchNorm2d, Relu, MaxPool2d]
  - Output Layer:1 x[FullyConnected(120)]
- activation Function
  - Input Layer: No Activation function
  - Hidden Layers: Relu
  - Output Layer: softmax

1.2 Loss Curves

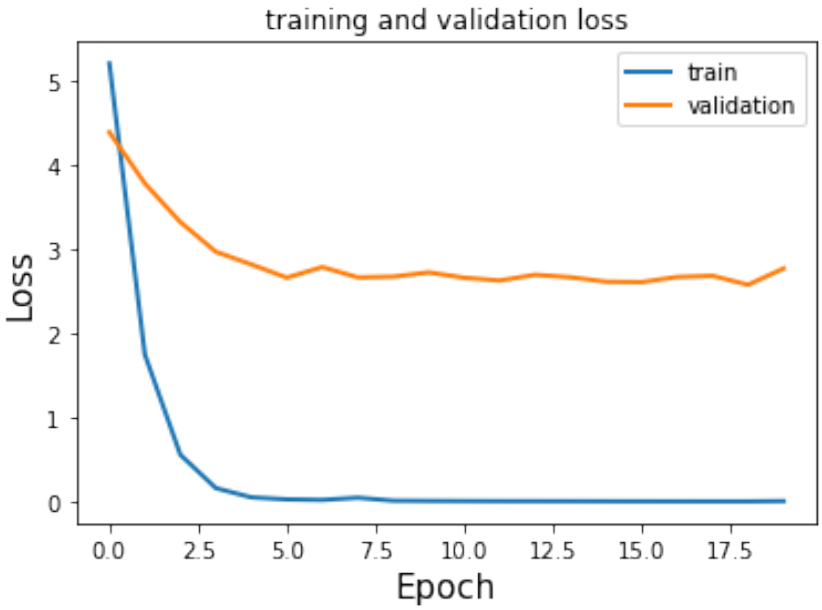

Figure 1: Task 1 loss curves

1.3 Accuracy Curves

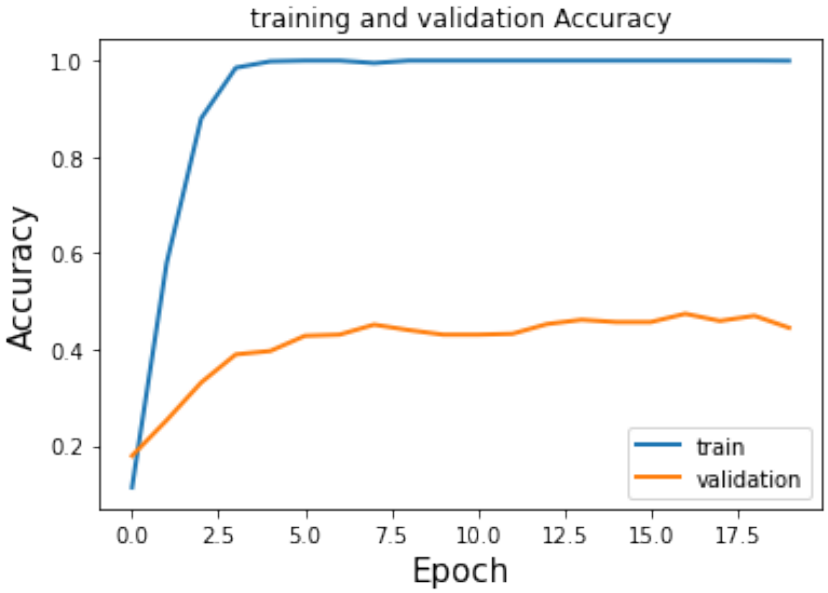

Figure 2: Task 1 accuracy curves

2 Task # 2

2.1 Model Parameters

- Learning rate = 0.001
- Optimizer = Adam
- Epochs = 10
- model Layer with neurons in each layer:
  - Input Layer: 3*218*178
  - Hidden Layers: 2X[Conv2d, BatchNorm2d, Relu, MaxPool2d]
  - Output Layer: 1 x[FullyConnected(10)]. Each image is passed from the network and then the difference of two output layers is taken and after that the sum of all 10 values in the difference vector is computed.
- activation Function
  - Input Layer: No Activation function
  - Hidden Layers: Relu
  - Output Layer: Sigmoid. The final output after difference and summation, which is a single value, is subjected to sigmoid.

2.2 Loss Curves

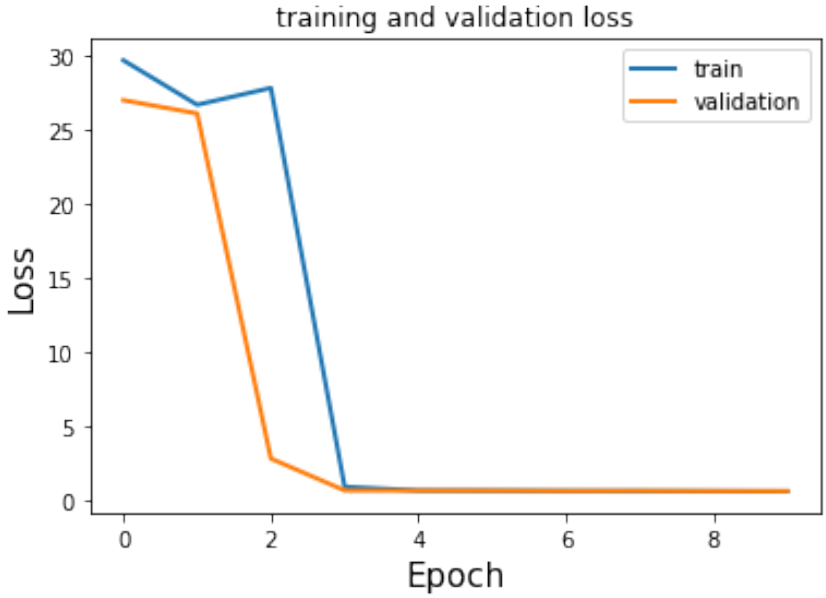

Figure 3: Task 2 loss curves

2.3 Accuracy Curves

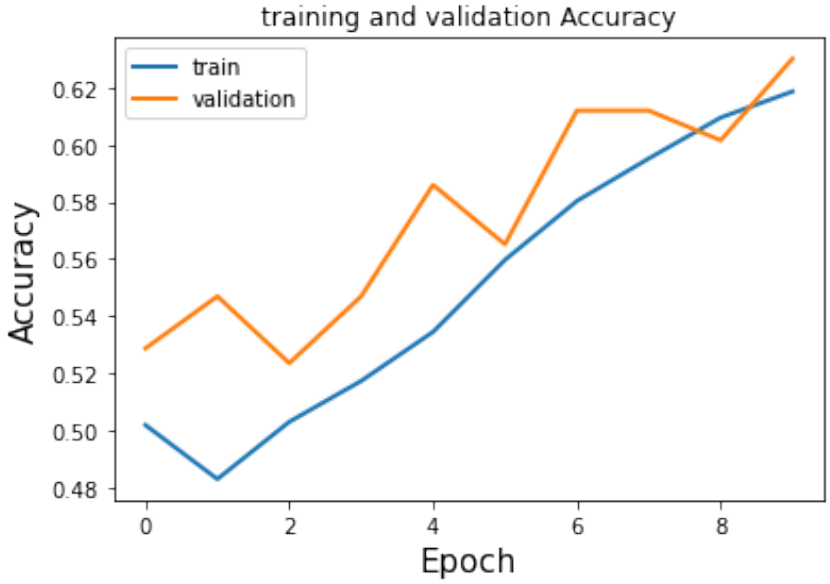

Figure 4: Task 2 accuracy curves

3 Task # 3

3.1 Model Parameters

- Learning rate = 0.0001
- Optimizer = Adam
- Epochs = 10

- model Layer with neurons in each layer:
  - Input Layer: 3*218*178
  - Hidden Layers: 3X[Conv2d, BatchNorm2d, Relu, MaxPool2d]
  - Output Layer: 1 x[FullyConnected(7)]
- activation Function
  - Input Layer: No Activation function
  - 4 Hidden Layer: Relu
  - Output Layer: sigmoid on each neuron.

3.2 Loss Curves

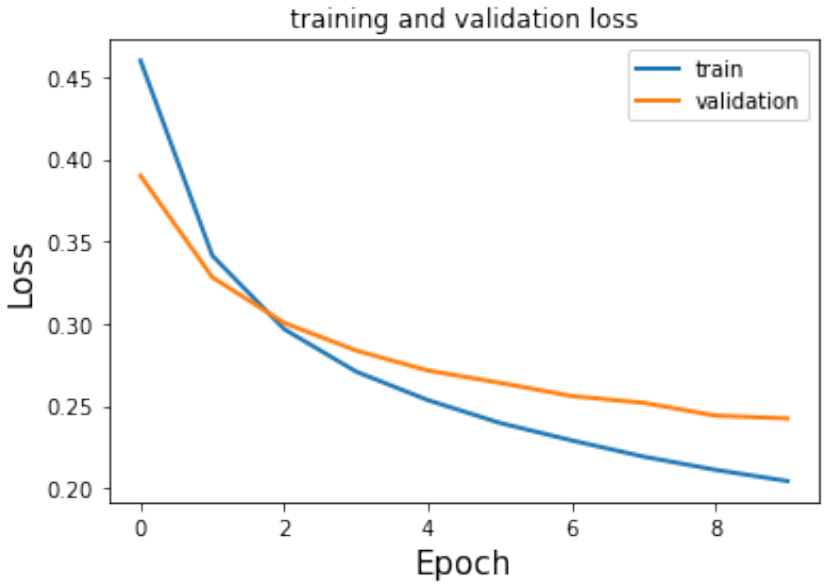

Figure 5: Task 3 loss curves

3.3 Accuracy Curves

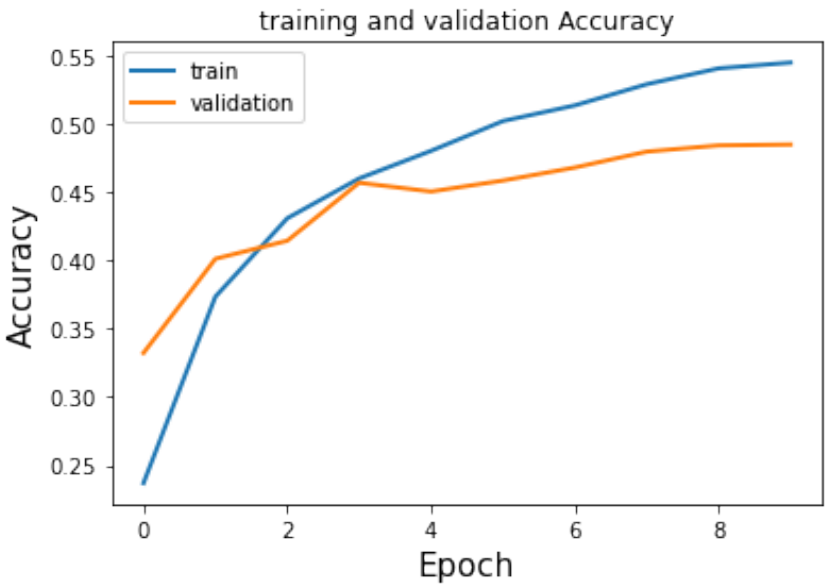

Figure 6: Task 3 accuracy curves

This model took 14+ hours to train and unfortunately I saved validation accuracy in both train and validation accuracy variables, hence the curves overlap.

4 Experiments

4.1 Task 1

The model architecture is kept same and only learning rate and epochs are changes

- Learning rate 0.001, epoch 1
  - train Accuracy: 10.7%
  - Validation Accuracy: 12.3%
  - test Accuracy: 39.11%
- Learning rate 0.001, epoch 10
  - train Accuracy: 100%
  - Validation Accuracy: 31.7%
  - test Accuracy: 34.11%
- Learning rate 0.01, epoch 10
  - train Accuracy: 91.6%
  - Validation Accuracy: 19.3%
  - test Accuracy: 21.2%
- Learning rate 0.0001, epoch 20
  - train Accuracy: 99.96%
  - Validation Accuracy: 44.46%
  - test Accuracy: 41.96%
- Learning rate 0.0001 (with decay), epoch 20

  - train Accuracy: 100%
  - Validation Accuracy: 42.38%
  - test Accuracy: 41.96%

It is to be noted that loss curve is showing a decreasing trend and accuracy curve is showing an increasing trend as shown above. Therefore, increasing number of epochs will improve the accuracy but because of time limitations results is shown for max 10 epochs.

4.2 Task 2

The model architecture is kept same and only learning rate and epochs are changes.

- Learning rate 0.0001, epoch 1
  - train Accuracy: 61.41%
  - Validation Accuracy: 59.63%
  - test Accuracy: 61.37%
- Learning rate 0.0001, epoch 5
  - train Accuracy: 65.25%
  - Validation Accuracy: 64.06%
  - test Accuracy: 63%
- Learning rate 0.0001, epoch 10
  - train Accuracy: 68.63%
  - Validation Accuracy: 65.10%
  - test Accuracy: 61.89%
- Learning rate 0.001, epoch 10
  - train Accuracy: 61.86%
  - Validation Accuracy: 63.02%
  - test Accuracy: 60.07%

4.3 Task 3

- Learning rate 0.0001, epoch 1
  - train Accuracy: 42.20%
  - Validation Accuracy: 45.40%
  - test Accuracy: 52.60%
- Learning rate 0.00001, epoch 10
  - train Accuracy: 54.48%
  - Validation Accuracy: 48.48%
  - test Accuracy: 52.33%