FOUNDATIONS OF MACHINE LEARNING

HOMEWORK 2 REPORT

Prepared by:

Name	Roll No.	Program
Arijeet	23M0742	MTech CSE
A Asish	23M0759	MTech CSE

1. Simple Dataset

Seed = 100

Table Values: (valid acc, valid loss)

Momentum = 0

Epochs / Learning Rate	0.1	0.01	0.001
50	0.98, 0.15	0.97, 0.34	0.72, 0.88
100	0.99, 0.047	0.98, 0.16	0.85, 0.61
200	0.99, 0.047	0.98, 0.16	0.95, 0.4

Best Model: Learning rate- 0.1, epochs- 100
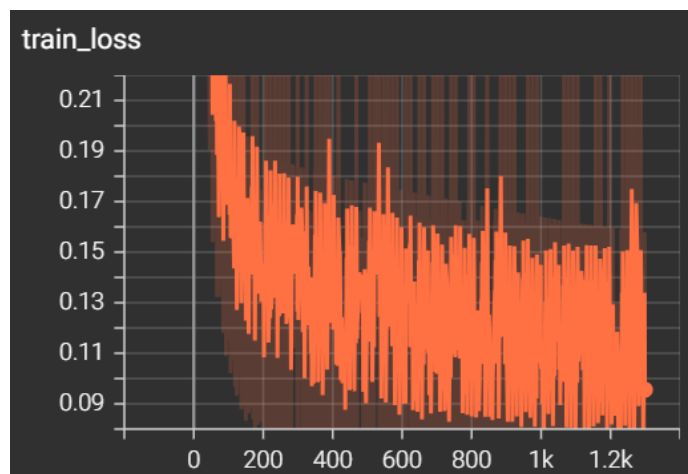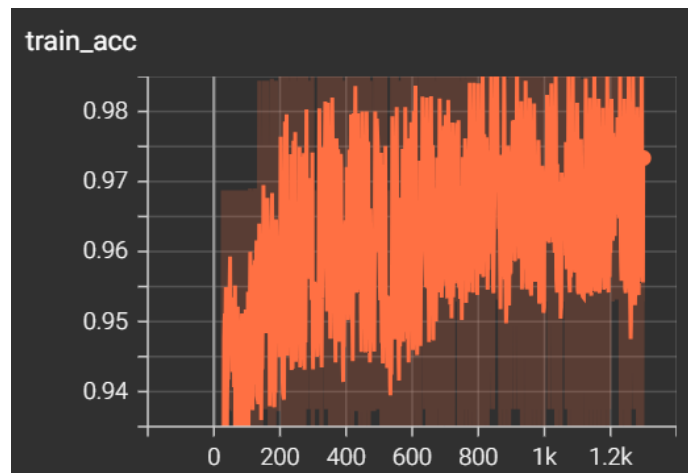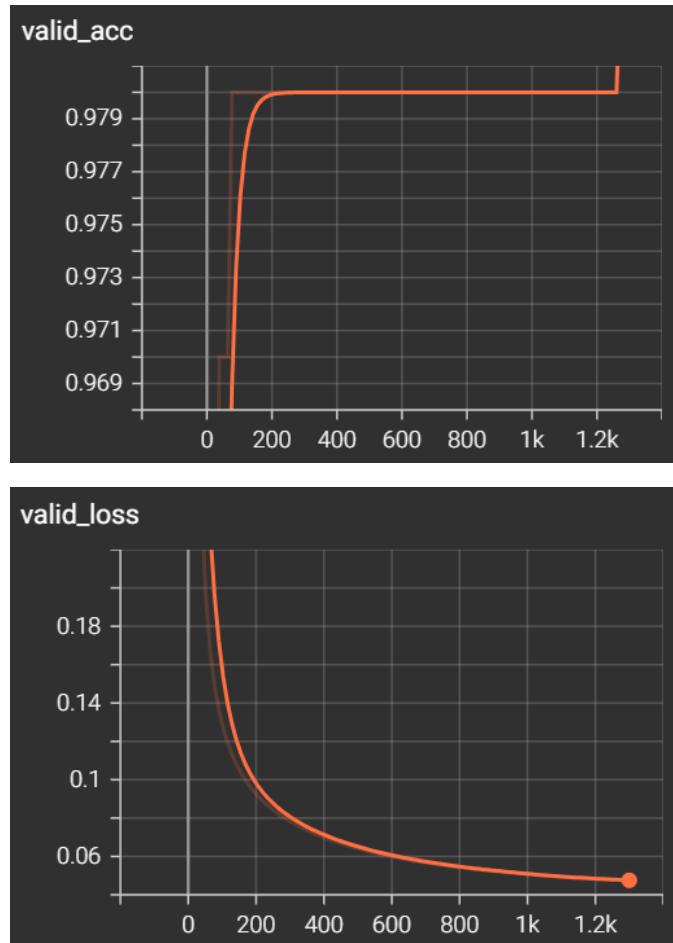
valid_acc

valid_loss

Various values of hyperparameters affected the accuracy and loss values in different ways.

- Learning rate: Having a low learning rate was observed to have given less accuracy as well as high loss. This is because the gradient descent takes very small steps due to small learning rate. Whereas a reasonable learning rate of 0.1 gave us good accuracy and lower loss value.

- Number of epochs: We observed that the model needed a moderate number of epochs to converge, to be precise we needed 100 epochs. Very low number of epochs were insufficient for the model to converge, whereas high number of epochs were unnecessary.

So, we observed that the hyperparameters that gave us the best results are learning rate = 0.1, number of epochs = 100. The validation accuracy we were able to obtain was 99% with a validation loss of 0.047.
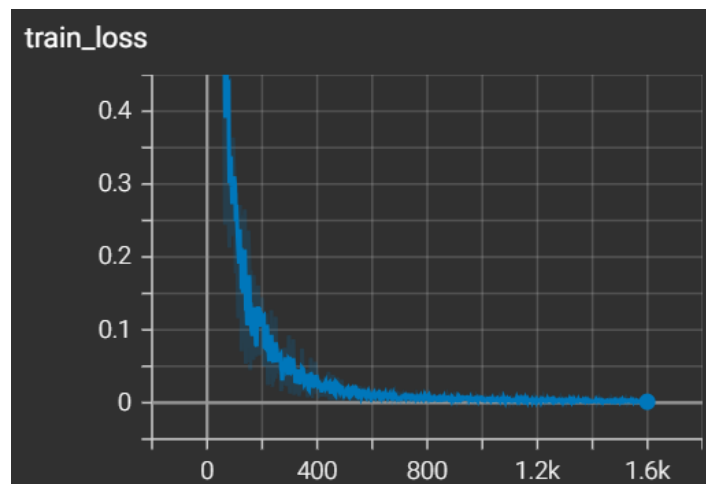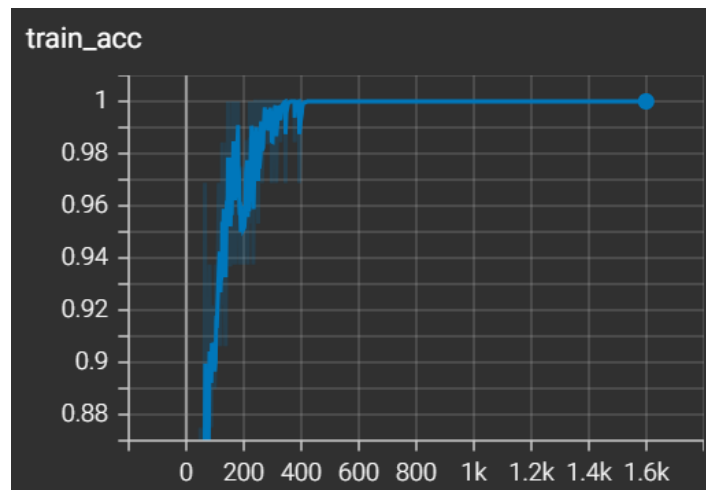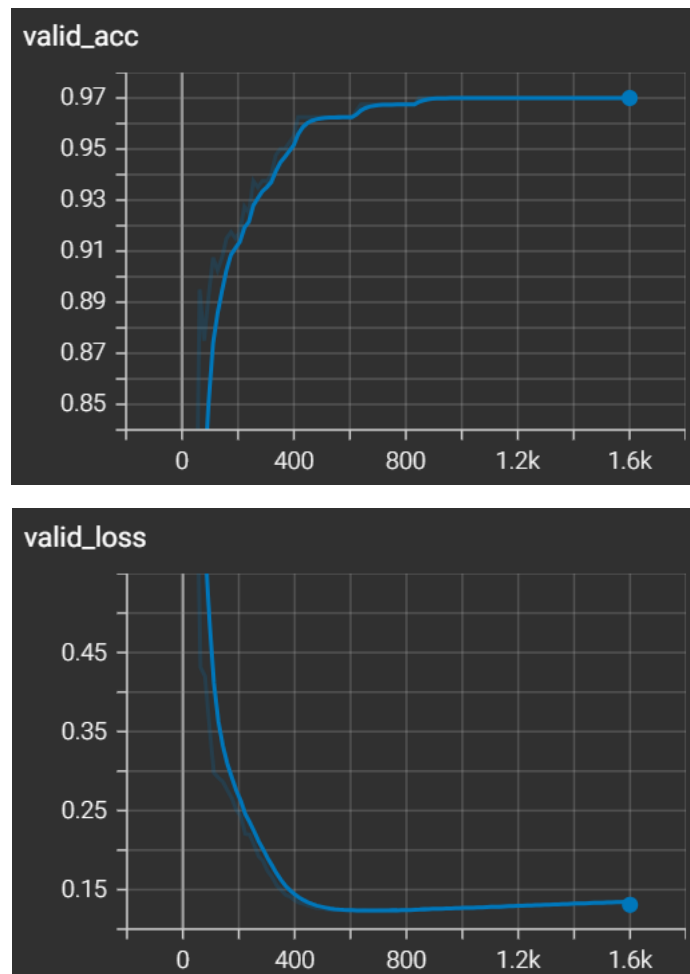
2. Digits Dataset

Seed = 100

Table Values: (valid acc, valid loss)

Momentum = 0

Epochs / Learning Rate	0.1	0.05	0.001
50	0.95, 0.18	0.96, 0.12	0.72, 1.21
100	0.95, 0.18	0.97, 0.12	0.85, 0.59
200	0.95, 0.18	0.97, 0.12	0.91, 0.32

Best Model: Learning rate - 0.05, epochs – 100

valid_acc

valid_loss

Various values of hyperparameters affected the accuracy and loss values in different ways.

- Learning rate: Having a low learning rate was observed to have given less accuracy as well as high loss. This is because the gradient descent takes very small steps due to small learning rate. Whereas a reasonable learning rate of 0.05 gave us good accuracy and lower loss value. Having high learning rate also resulted in less accuracy and more loss, because the algorithm was taking large steps and didn't converge properly at minima.

- Number of epochs: We observed that the model needed a moderate number of epochs to converge, to be precise we needed 100 epochs. Very low number of epochs were insufficient for the model to converge, whereas high number of epochs were unnecessary.

So, we observed that the hyperparameters that gave us the best results are learning rate= 0.05, number of epochs= 100. The validation accuracy we were able to obtain was 97% with a validation loss of 0.125.