MSBD 5012 HA5 report

Student Name: Zuo Yifan
Student ID: 20876522
Assignment #: Assignment HA5
Student Email: yzuoah@connect.ust.hk
Course Name: MSBD5012

- Baseline:
  - banch_size=32
    lr= le-5
    dropout=0.1

Best module Loss	Accuracy
Training Loss: 0.063 Validation Loss: 0.551	0.8

- Alter number of batch size experiment:
  - Increasing batch size from 32 to 48:

Best module Loss	Accuracy
Training Loss: 0.007 Validation Loss: 0.455	0.86

    Compare to the baseline result, we can see that the training loss and validation loss decreases and the accuracy increase a little which is due to the large batch size makes the convergence more efficient with less randomness.

  - Decreasing batch size from 32 to 16:

Best module Loss	Accuracy
Epoch 1 / 10 Training Loss: 0.677 Validation Loss: 0.656	0.73

    Compare to the baseline result, we can see that the loss increase for both tranining and validation and the accuracy decreases, because the less batch size we have, the more randomness will envolve during traning.

- Alter dropout experiment:
  - Increase dropout rate:

    We increase dropout rate from 0.1 to 0.8.

Loss	Accuracy
Epoch 2 / 10 Training Loss: 0.977 Validation Loss: 0.356	0.46

`Compare to the baseline result, we can see that traning loss increase a lot and validation loss decrease a bit and accuracy decreases , this is because the module is underfitting the datat since the dropout rate is too high too much parameters are dropout during training.`

- Decrease dropout rate:

  We increase dropout rate from 0.1 to 0.

Loss	Accuracy
Epoch 5 / 10 Training Loss: 0.012 Validation Loss: 0.756	0.49

`Compare to the baseline result, we can see that traning loss decrease a lot and validation loss increase a lot and accuracy decreases , this is because the module is overfitting the datat since the dropout rate is zero, no parameters are dropout during training.`

- Alter learning rate experiment:
  - Increasing learning rate:

    Learning rate from 10^{-5} to 10^{-3}

Best module Loss	Accuracy
Epoch 9 / 10 Evaluating... SAVING MODEL Training Loss: 0.695 Validation Loss: 0.693	0.5

`Compare to the baseline result, we can see that the loss decreases and the accuracy decreases a lot, becrease in gradient descent process, we rising learning rate too much which leads to the step moved in the opposite direction of the derivative too much to miss the minima which leads to slow learning and low accuracy comparing with same amount of learning.`

  - Increase learning rate:

    Learning rate from 10^{-5} to 10^{-4}

Best module Loss	Accuracy
Epoch 1 / 10 Evaluating... SAVING MODEL Training Loss: 0.693 Validation Loss: 0.693	0.51

`The same comment as the previous experiment`

- Decreasing learning rate:

Learning rate from 10^{-5} to 10^{-6}

Best module Loss	Accuracy
Epoch 9 / 10 Evaluating... SAVING MODEL Training Loss: 0.687 Validation Loss: 0.687	0.54

`Compare to the baseline result, we can see that the loss decreases and the accuracy`
`decreases a lot, since decreasing learning rate means shrinking the step moved in the`
`opposite direction of the derivative which leads to slow learning and low accuracy`
`comparing with same amount of learning.`

- Alter maximum number of epochs experiment:

  - Increase the number of epochs from 10 to 15:

Best modukle Loss	Accuracy
Training Loss: 0.020 Validation Loss: 0.495 Epoch 2 / 15	0.82

`Compare to the baseline result, we see the traning loss and the validation loss both`
`decreases, and the accuracy increases, becuase increasing number of epochs will rise`
`the possibility of finding a better solution`

  - Decrease the number epochs from 10 to 5:

Best module Loss	Accuracy
Training Loss: 0.015 Validation Loss: 0.606 7/5	0.79

`Compare to the baseline result, we see the traning loss decrease and the validation`
`loss increases, and the accuracy decrease a little bit, becuase decrease number of`
`epochs will drop the possibility of finding a better solution meaning that the`
`solution we find is really a local best solution`

  - Decrease the number of epochs from 10 to 1:

Best module Loss	Accuracy
Training Loss: 0.012 Validation Loss: 0.326 1/1	0.86

Compare to the baseline result, we see that the traning loss and validation loss both decrease, and accuray increases I guess this is just the lucky experiment that we find a relatively good solution in a local space among the global space.