

AssignmentReport-Group199

February 23, 2024

1 Assignment 2 Report

2 Task 1

2.1 task 1a)

$i \rightarrow j \rightarrow k$

TDT 4265 Assignment 2

Task 1

$$a) \quad \frac{\partial \mathcal{L}}{\partial w_{ji}} = \frac{\partial \mathcal{L}}{\partial z_j} \cdot \frac{\partial z_j}{\partial w_{ji}}$$

$$\frac{\partial z_j}{\partial w_{ji}} = \frac{\partial}{\partial w_{ji}} (\dots + x_i w_{ij} + x_{i+1} w_{i+1j} + \dots)$$
$$= x_i$$

Therefore:

$$w_{ji} = w_{ji} - \alpha \delta_j x_i$$

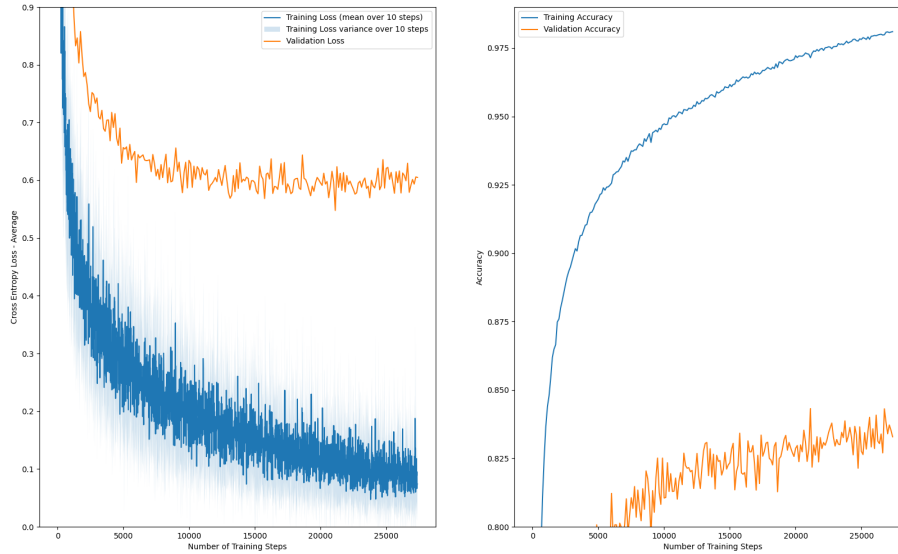
$$\text{where } \delta_j = \frac{\partial \mathcal{L}}{\partial z_j}$$

this will be the sum of all output node errors it affects, scaled by the activation function

$$\delta_j = \frac{\partial \mathcal{L}}{\partial z_j} = \underbrace{f'(z_j)}_{\text{activation function}} \underbrace{\sum_k w_{kj} \delta_k}_{\text{output error effect}}$$

3 Task 2

3.1 Task 2c)



Training stopped after 44 epochs.

Final Train Cross Entropy Loss: 0.08452529520250651

Final Validation Cross Entropy Loss: 0.6049188886493303

Train accuracy: 0.981

Validation accuracy: 0.8329

3.2 Task 2d)

$785 \times 64 + 64 \times 10 = 50\,880$

4 Task 3

After adding improved weights

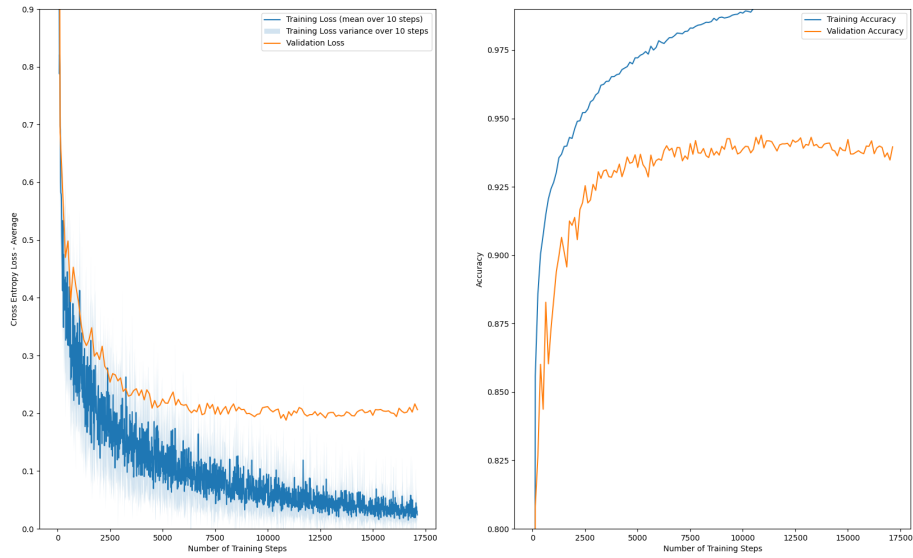
Training stopped after 28 epochs.

Final Train Cross Entropy Loss: 0.028842773450010798

Final Validation Cross Entropy Loss: 0.2061919544861595

Train accuracy: 0.9969

Validation accuracy: 0.9396



After improved sigmoid

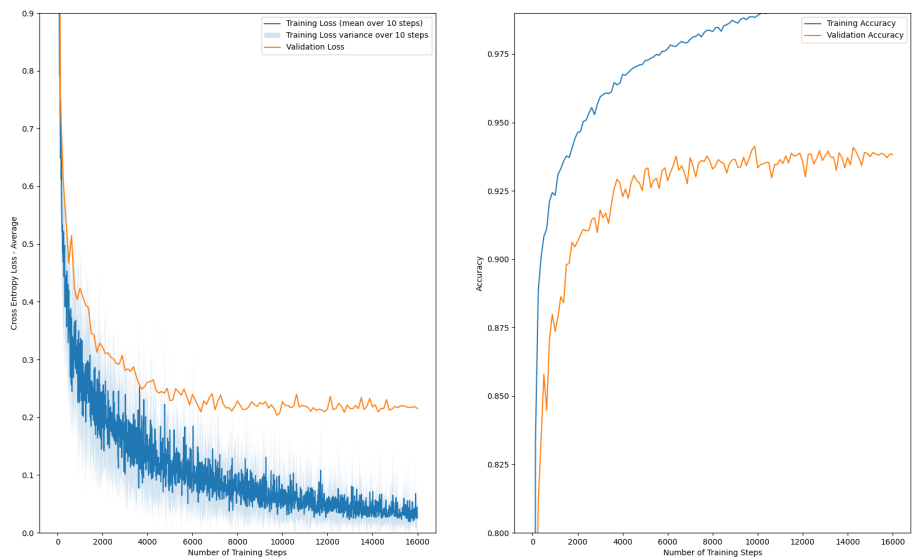
Training stopped after 26 epochs.

Final Train Cross Entropy Loss: 0.03162280740067008

Final Validation Cross Entropy Loss: 0.21526426614082944

Train accuracy: 0.99595

Validation accuracy: 0.9383



After momentum

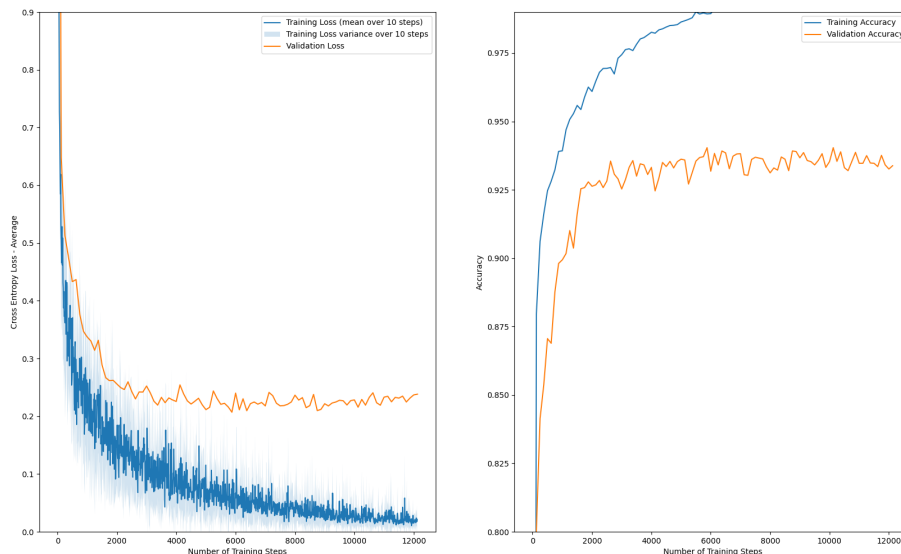
Training stopped after 20 epochs.

Final Train Cross Entropy Loss: 0.018526288809209552

Final Validation Cross Entropy Loss: 0.23824469455066005

Train accuracy: 0.99855

Validation accuracy: 0.9338



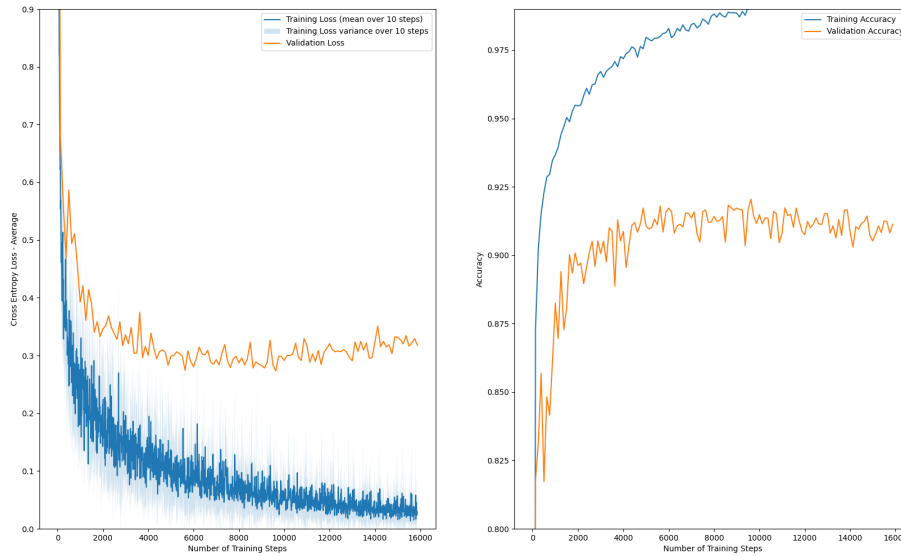
After adding the improved weights both the training and validation accuracy jumps up, however they both dip down again after improved sigmoid, then the train acc goes up again with momentum while validation keeps going down. These changes are very small however, and i dont think they are significant. If i did it wrong i feel like the change would be much more drastic.

Also, the convergence speed increases with each added trick. From the base model in task 2 the training time is decreased by more than half.

We do however see significant overfitting in all models, evident by the validation accuracy and loss flattening out, while the training acc keep indreasing and the training loss keep decreasing. I think a contributing factor to this is a too strict rule for when the program is allowed to stop training.

5 Task 4

5.1 Task 4a)

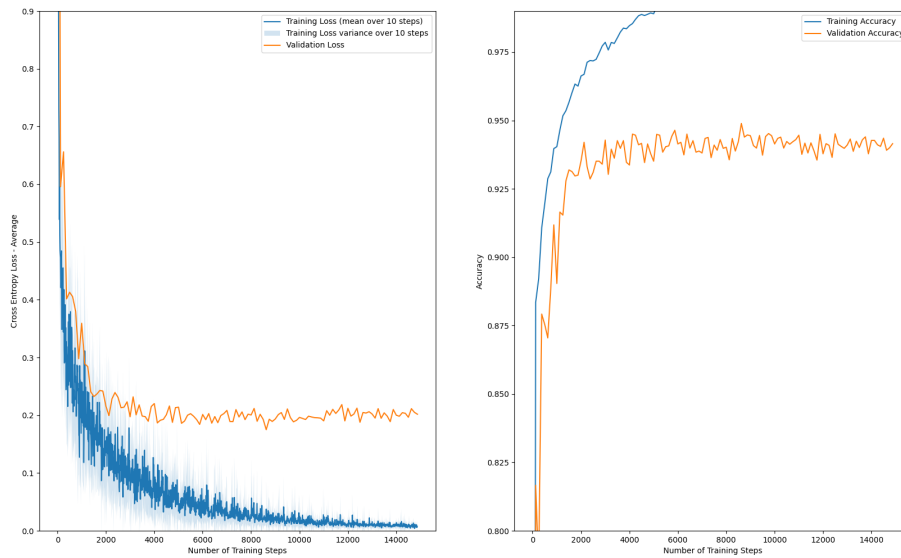


Train accuracy: 0.99645

Validation accuracy: 0.9113

This seems to be underfitting, as the accuracy plateaus earlier than what was possible with a higher amount of nodes in the hidden layer

5.2 Task 4b)

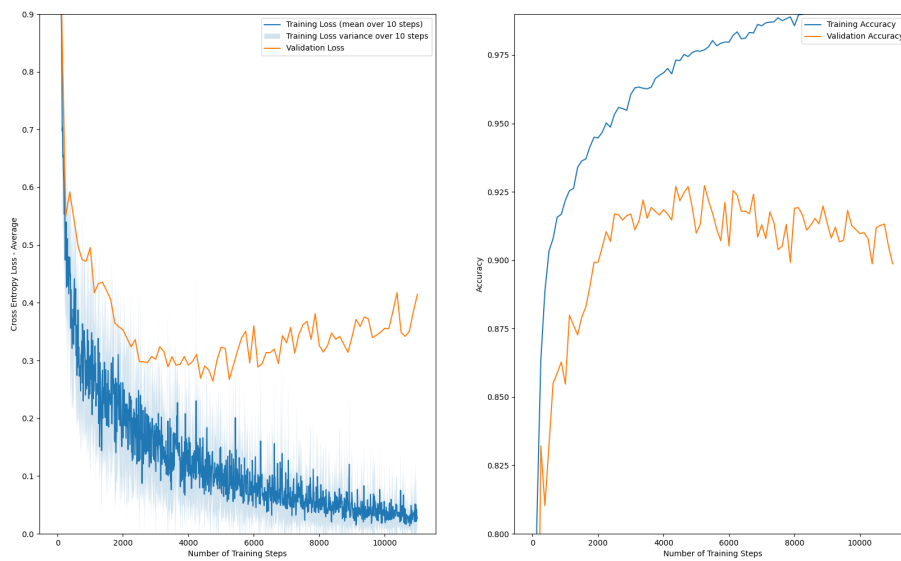


Train accuracy: 1.0

Validation accuracy: 0.9415

This is definitely overfitting, as a training accuracy of 1 is a sign of overfitting, not a perfect model.

5.3 Task 4d)



Training stopped after 18 epochs.

Final Train Cross Entropy Loss: 0.02569169268081729
Final Validation Cross Entropy Loss: 0.4148459155837418
Train accuracy: 0.9962
Validation accuracy: 0.8987

In task 3 there was $50,880 \times 2 = 101,760$ parameters, times two from task 2 because we also have a momentum for each weight. With an extra hidden layer this increased to $(785 \times 64 + 64 \times 64 + 64 \times 10) \times 2 = 109,952$ parameters

The overfitting seems to be even worse, as the validation accuracy dont just flatten out but actually decrease. I tried to decrease the epochs but the result was still a bit worse than the one layer. I assume this is because i have done something wrong, but i dont know what

5.4 Task 4e)

Unfortunately i couldnt get this working on my pc. it just froze up and the vsc crashed. For the next assignment i will be sure to figure how to connect to the cybele pc's and run the code on them, but i unfortunately dont have the time for that now.