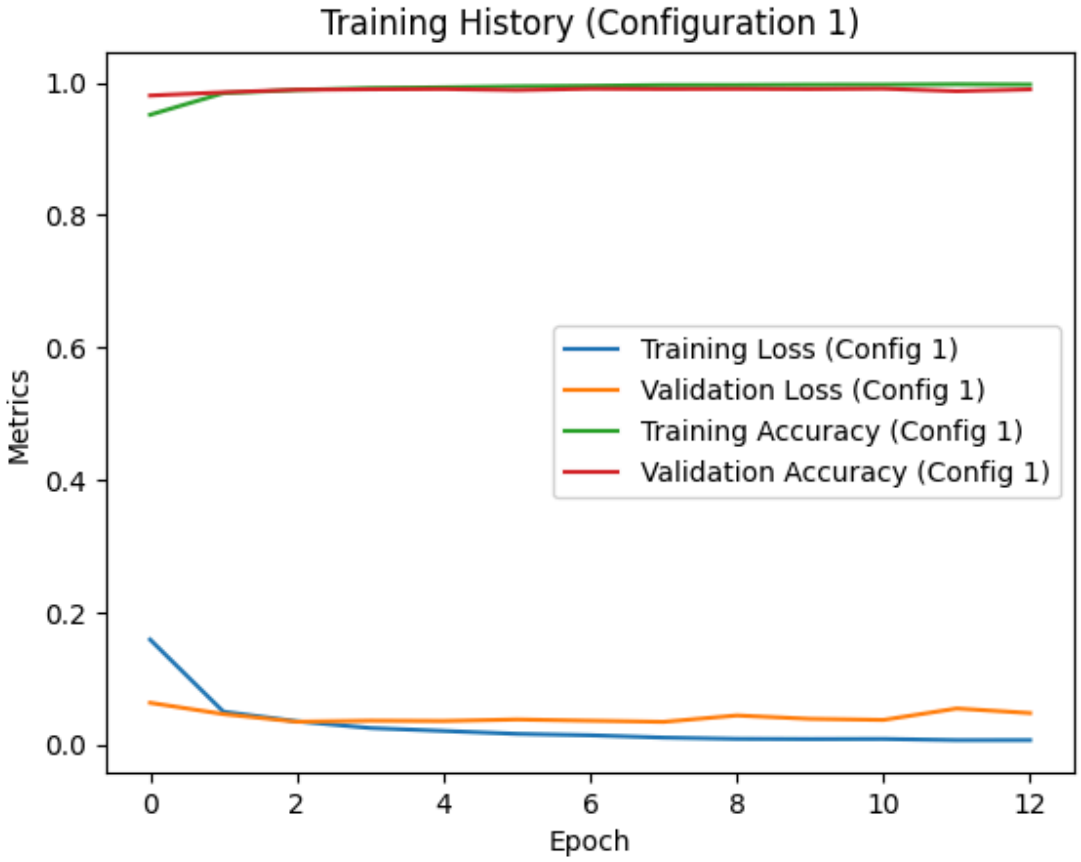
DAMA61 - 5th assignment

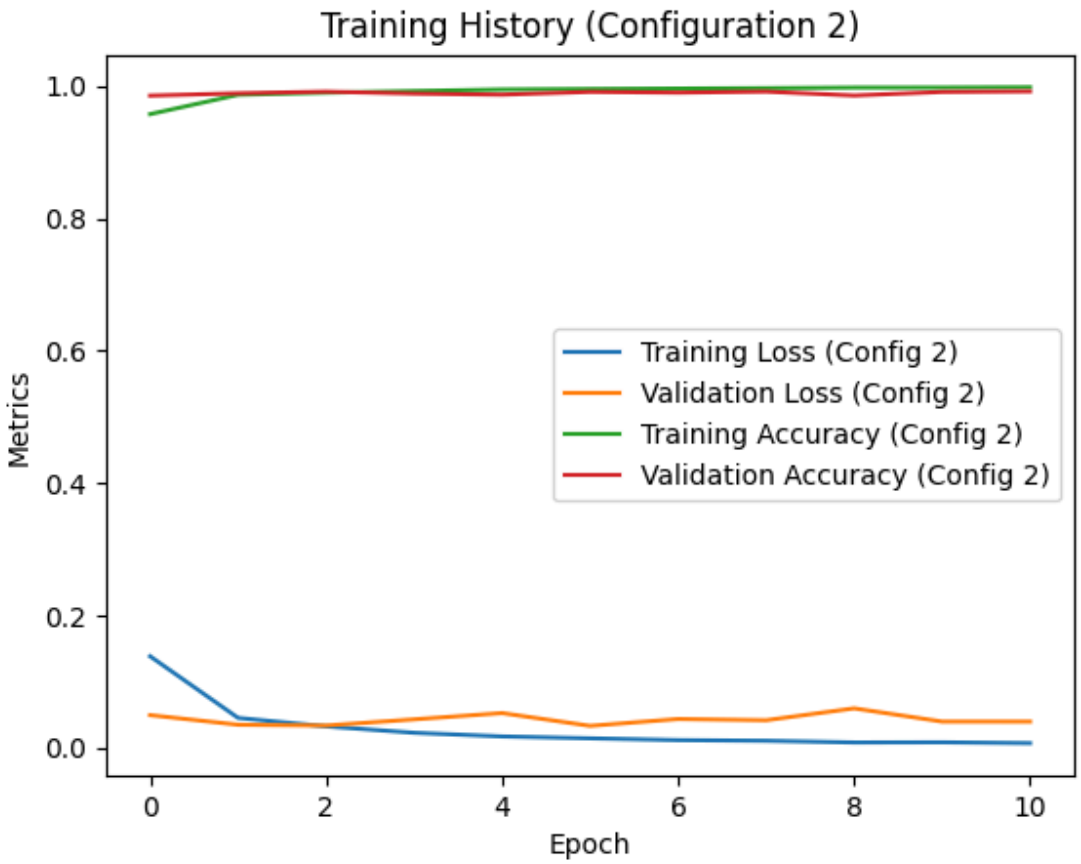
Exercise 1

1-6

All the instructions from the exercises were followed. The code produced the following plots:



DAMA61 - 5th assignment



Console output:
Test accuracy for model configuration 1: 0.9904000163078308
Test accuracy for model configuration 2: 0.9911999702453613

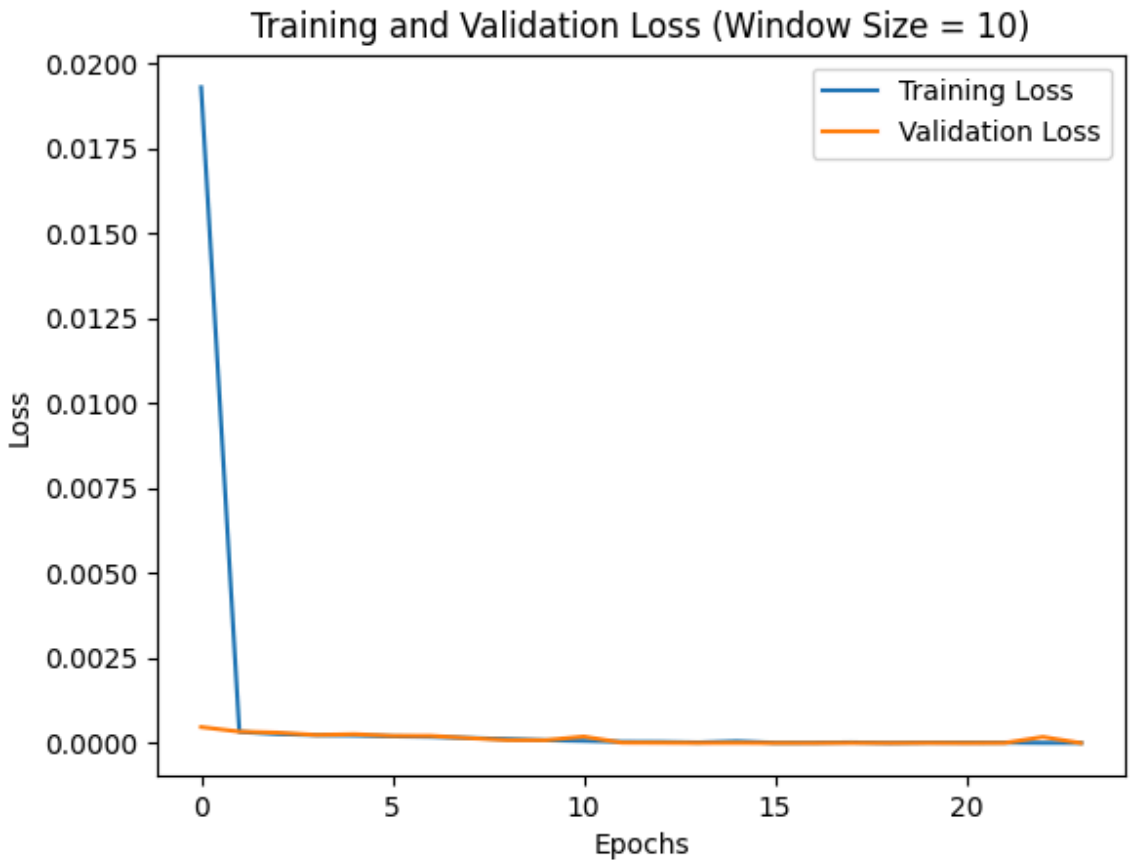
The models trained with both configurations achieved high test accuracies of approximately 99%, indicating that they generalize well to unseen data. Despite the observed trends of high training and validation accuracies, the models did not suffer from significant overfitting, as evidenced by their strong performance on the test set. These results suggest that the chosen architectures effectively captured the underlying patterns in the MNIST dataset.

DAMA61 - 5th assignment

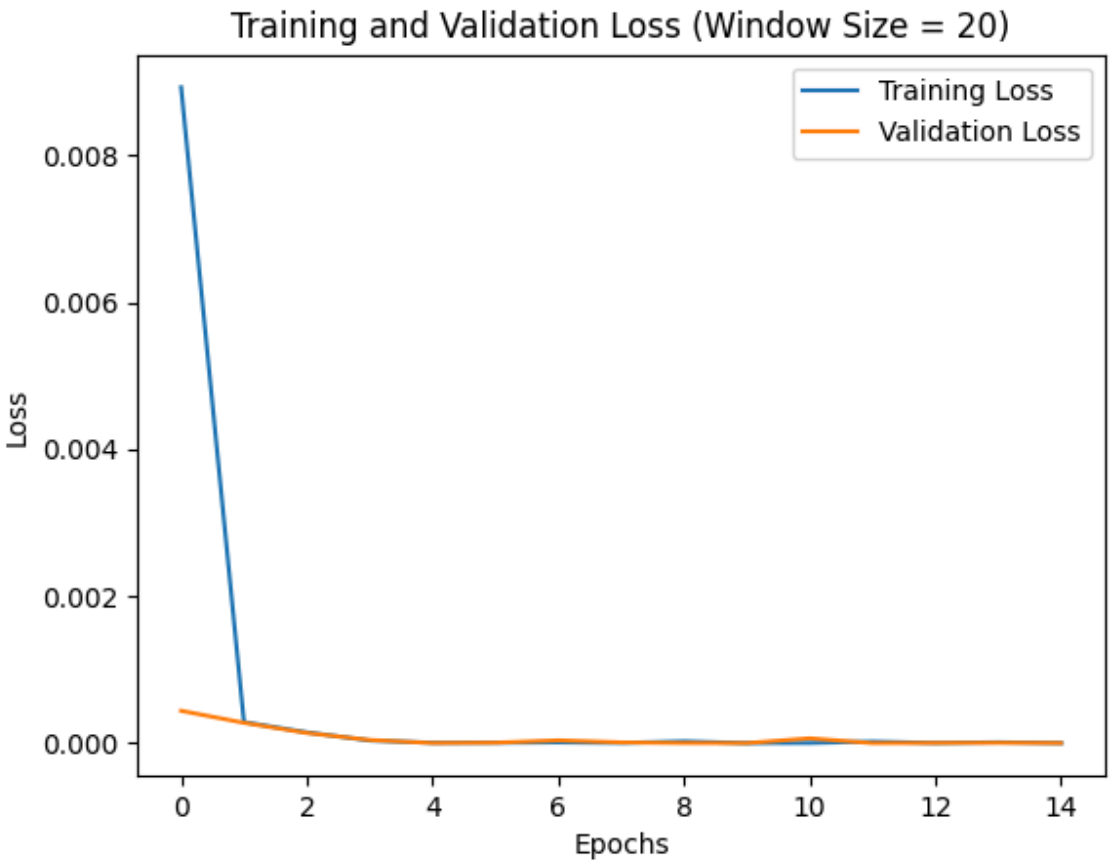
Exercise 2

1-7

All the instructions from the exercises were followed. The code produced the following plots:



DAMA61 - 5th assignment



Console output:
Test Loss (Window Size = 10): 0.00000103
Test Loss (Window Size = 20): 0.00000035

Both models exhibit a decrease in training loss, starting from around 0.019 for window size 10 and 0.009 for window size 20, dropping close to 0 within the first epoch. Despite both models achieving low training loss values, the model with a window size of 20 performs slightly better on the test set, as evidenced by its lower test loss compared to the model with a window size of 10. Increasing the window size from 10 to 20 results in a model that performs slightly better on the test set, indicating that a larger context window allows the model to capture more relevant temporal dependencies in the data. It is essential to keep in mind the test samples results, to know if the model overfits the data (which doesn't as we see on the results), but

even though the larger window does perform better, the increase in accuracy is small enough to make it not worth in cases that runtime is of importance.