Here is the base line option for our code.
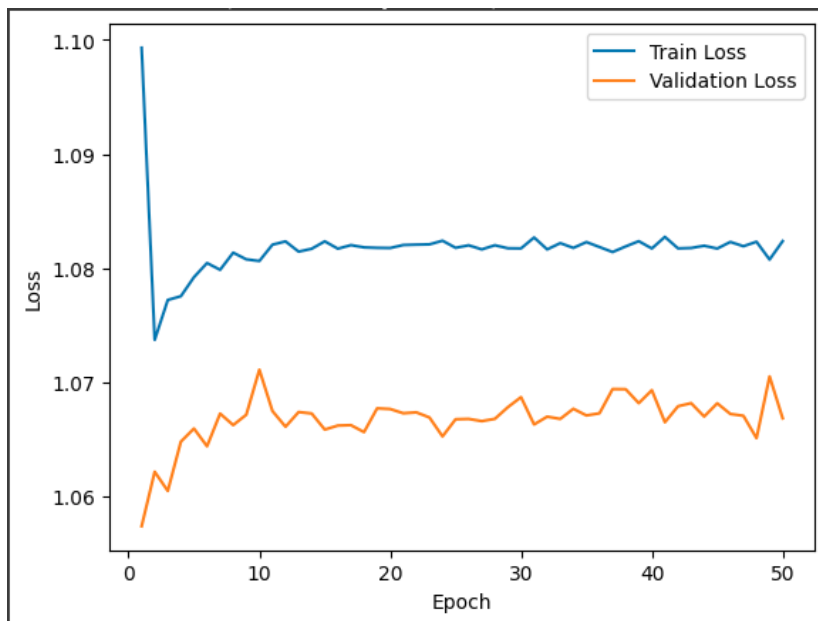
Normalization = True

Mini_batch_size_1

Model weight: Xavier

Learning Rate = 0.001

L2_lamba = 0.1

Optimizer: Momentum
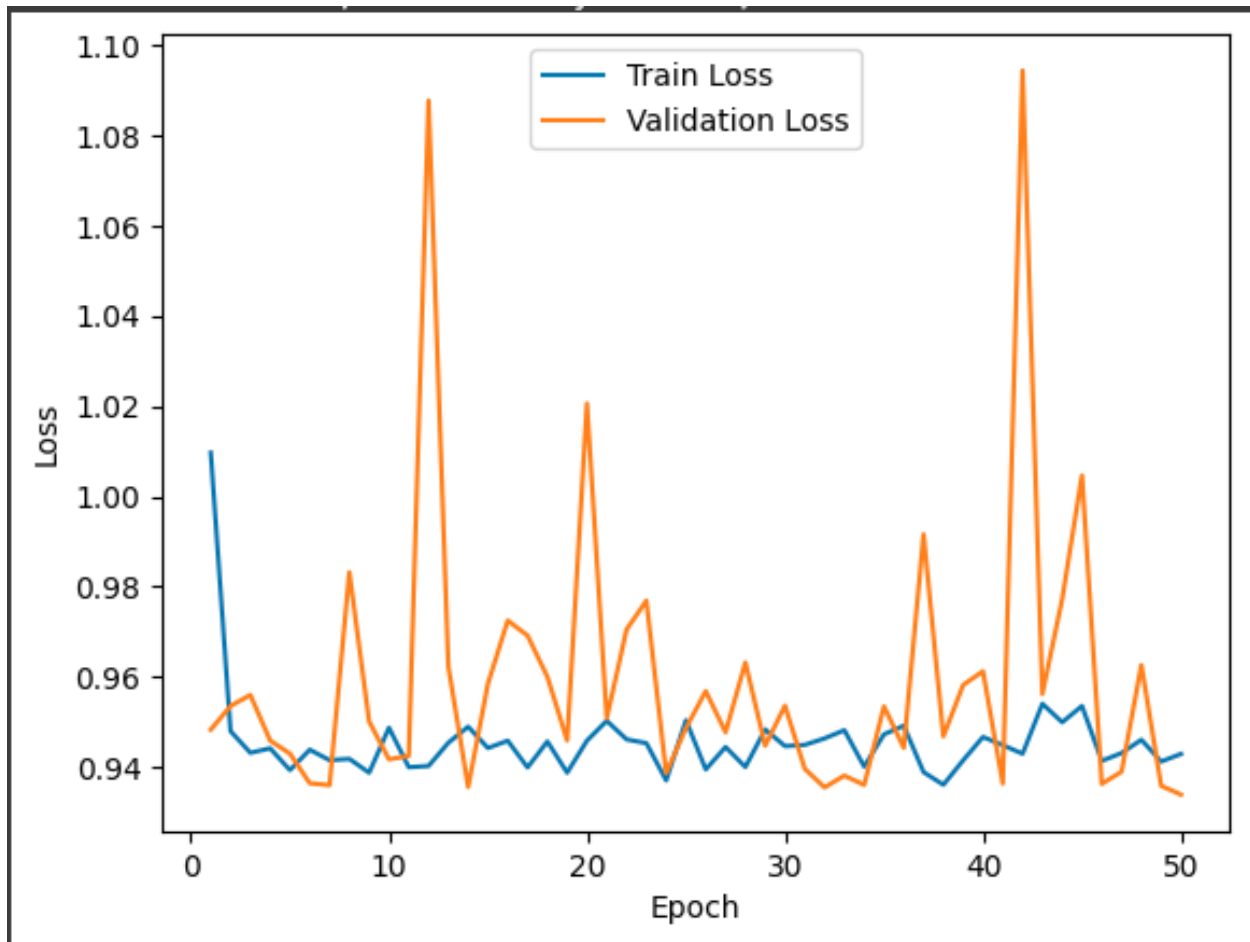
Epoch = 50

Normalization = False

Mini_batch_size_1

Model weight: Xavier

Learning Rate = 0.001

L2_lamba = 0.1

Optimizer: Momentum
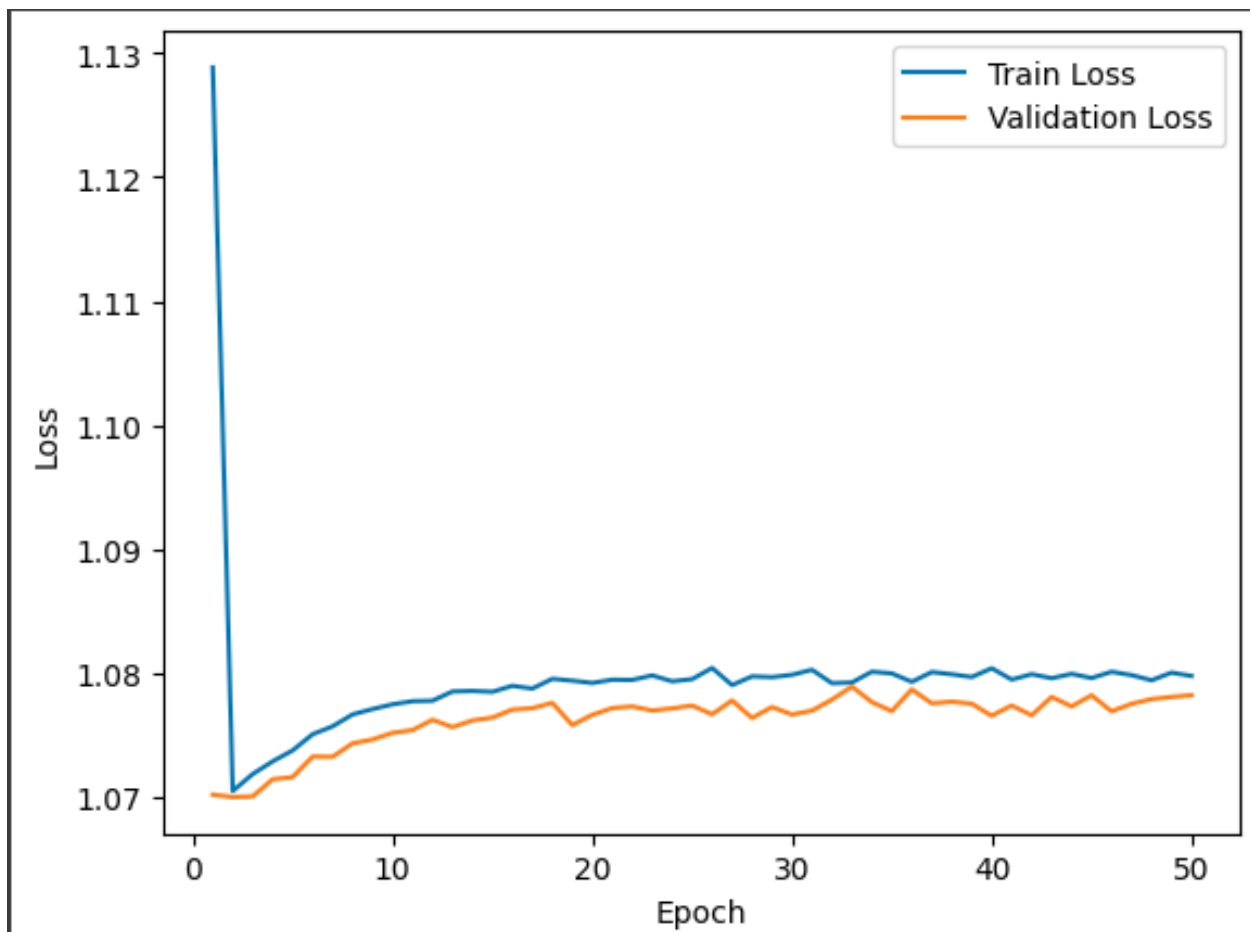
Epoch = 50

Normalization = True

Mini_batch_size_2

Model weight: Xavier

Learning Rate = 0.001

L2_lamba = 0.1

Optimizer: Momentum
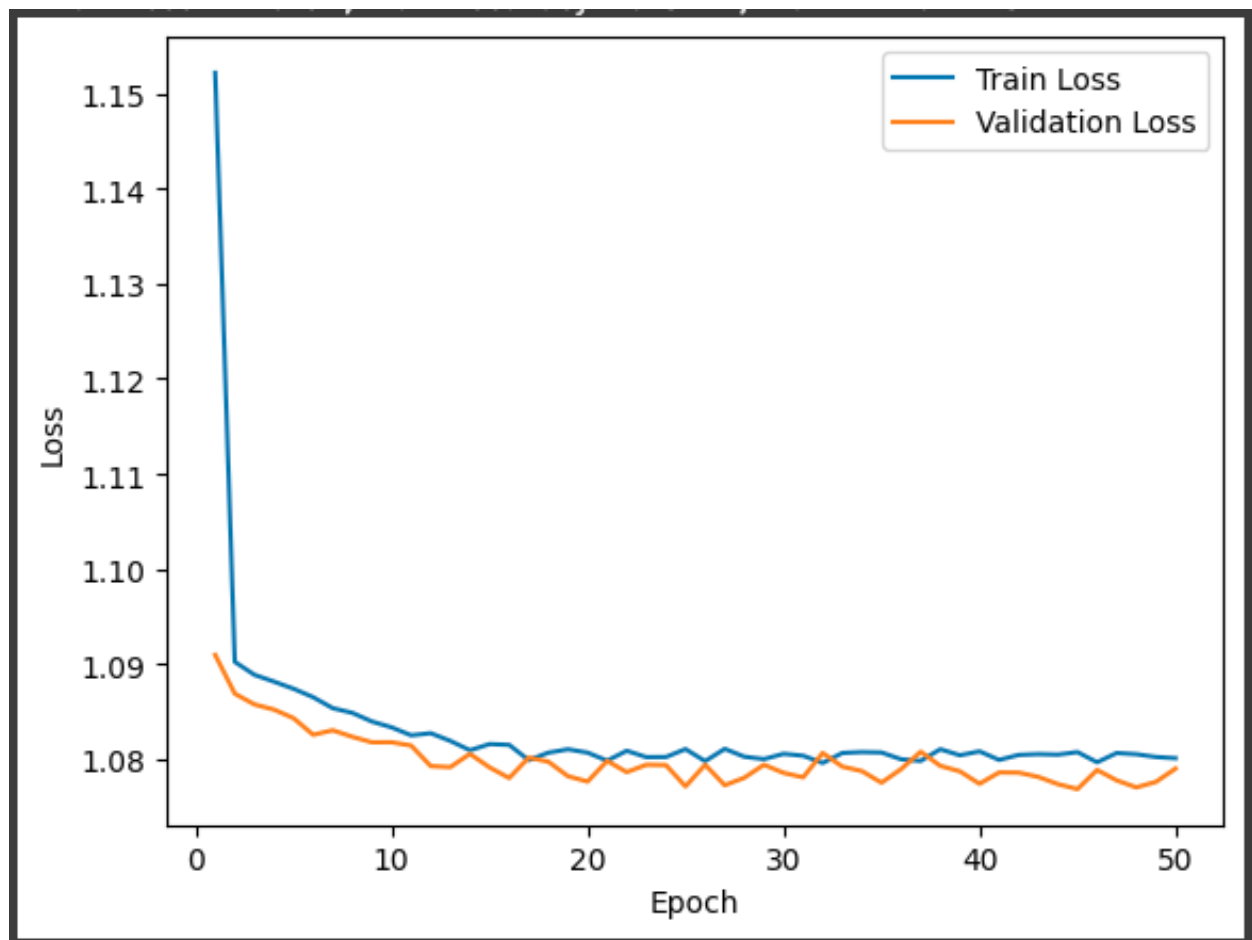
Epoch = 50

Normalization = True

Mini_batch_size_1

Model weight: Random

Learning Rate = 0.001

L2_lamba = 0.1

Optimizer: Momentum
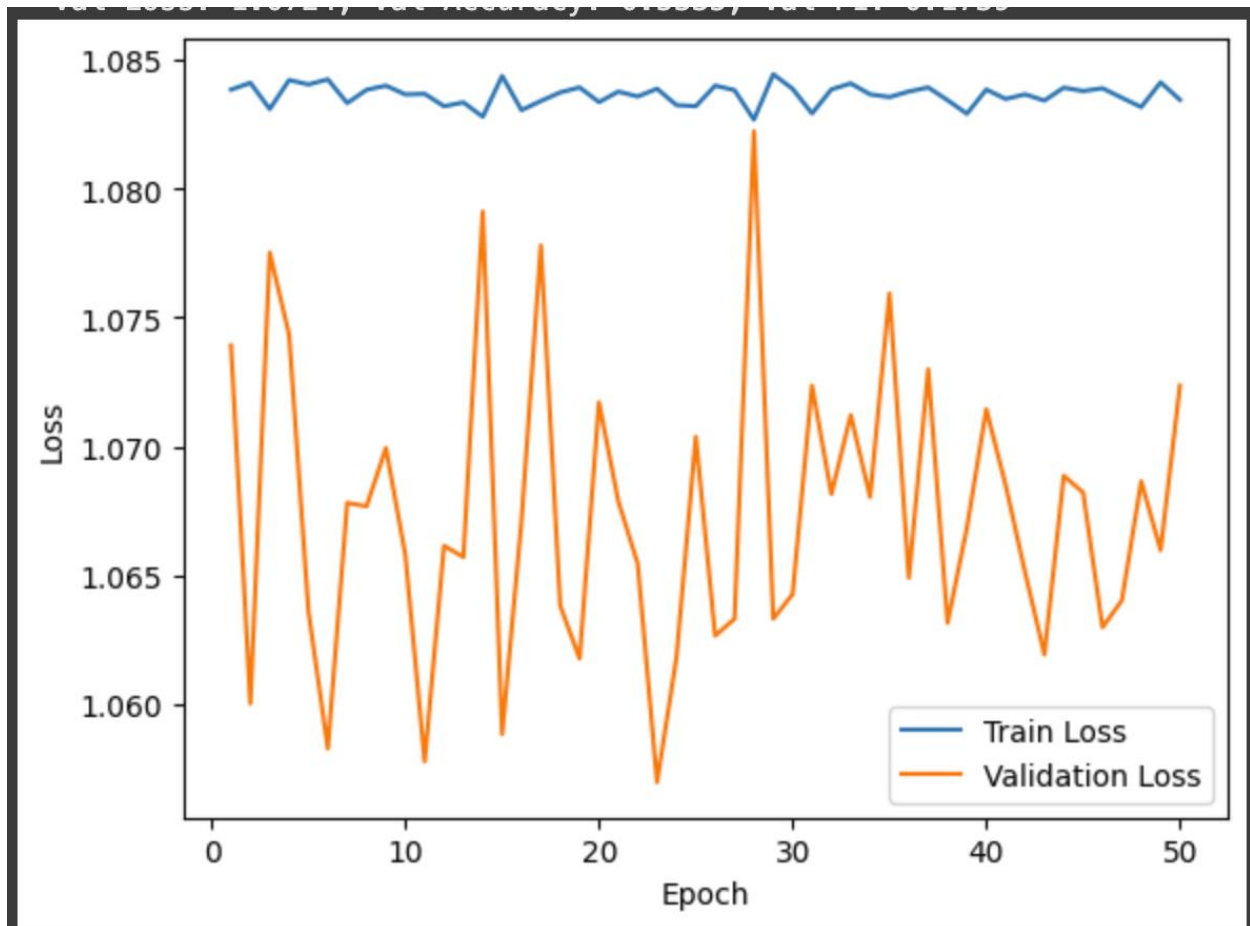
Epoch = 50

Normalization = True

Mini_batch_size_1

Model weight: Xavier

Learning Rate = 0.01

L2_lamba = 0.1

Optimizer: Momentum
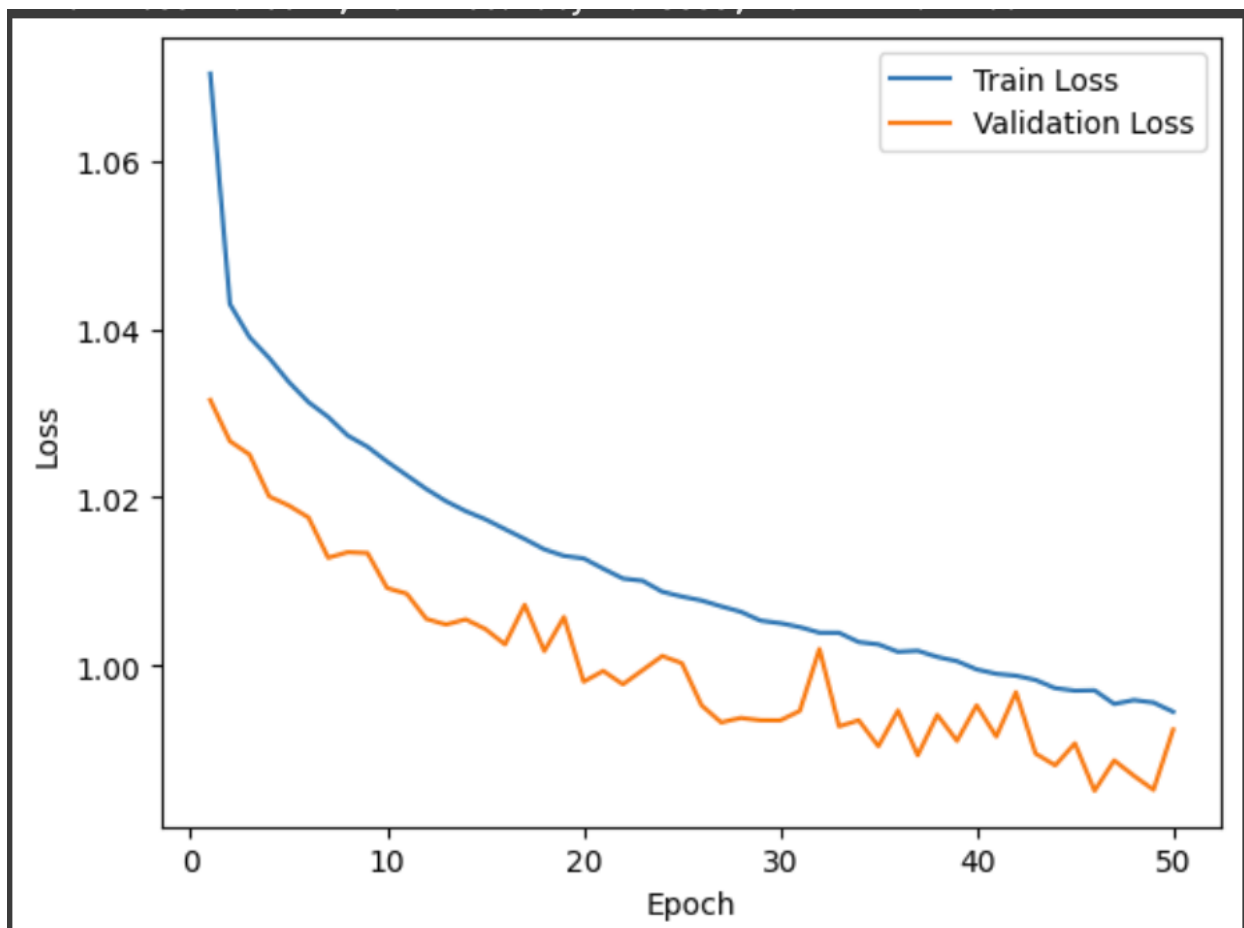
Epoch = 50

Normalization = True

Mini_batch_size_1

Model weight: Xavier

Learning Rate = 0.001

L2_lamba = 0

Optimizer: Momentum
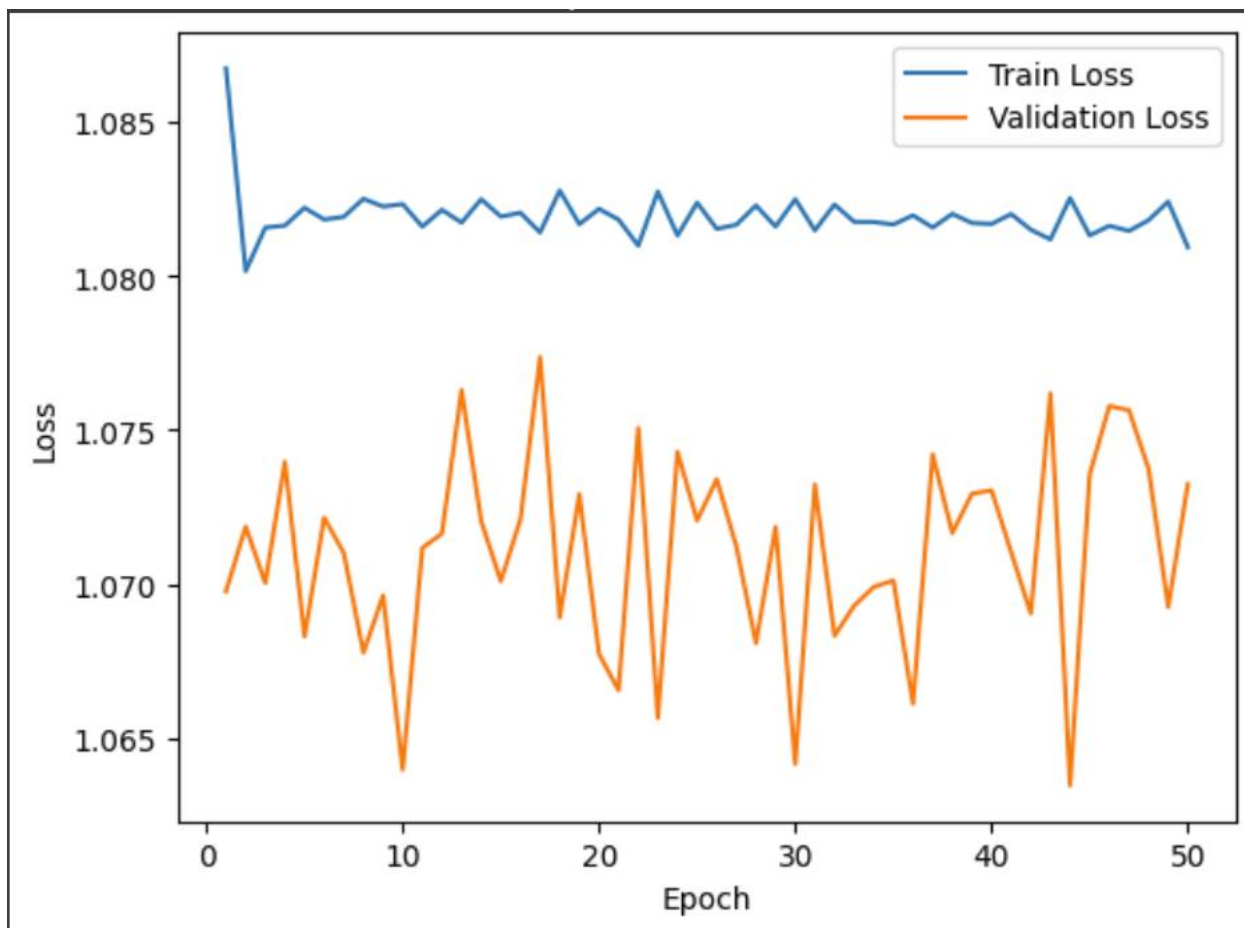
Epoch = 50

Normalization = True

Mini_batch_size_1

Model weight: Xavier

Learning Rate = 0.001

L2_lamba = 0.1

Optimizer: Adam
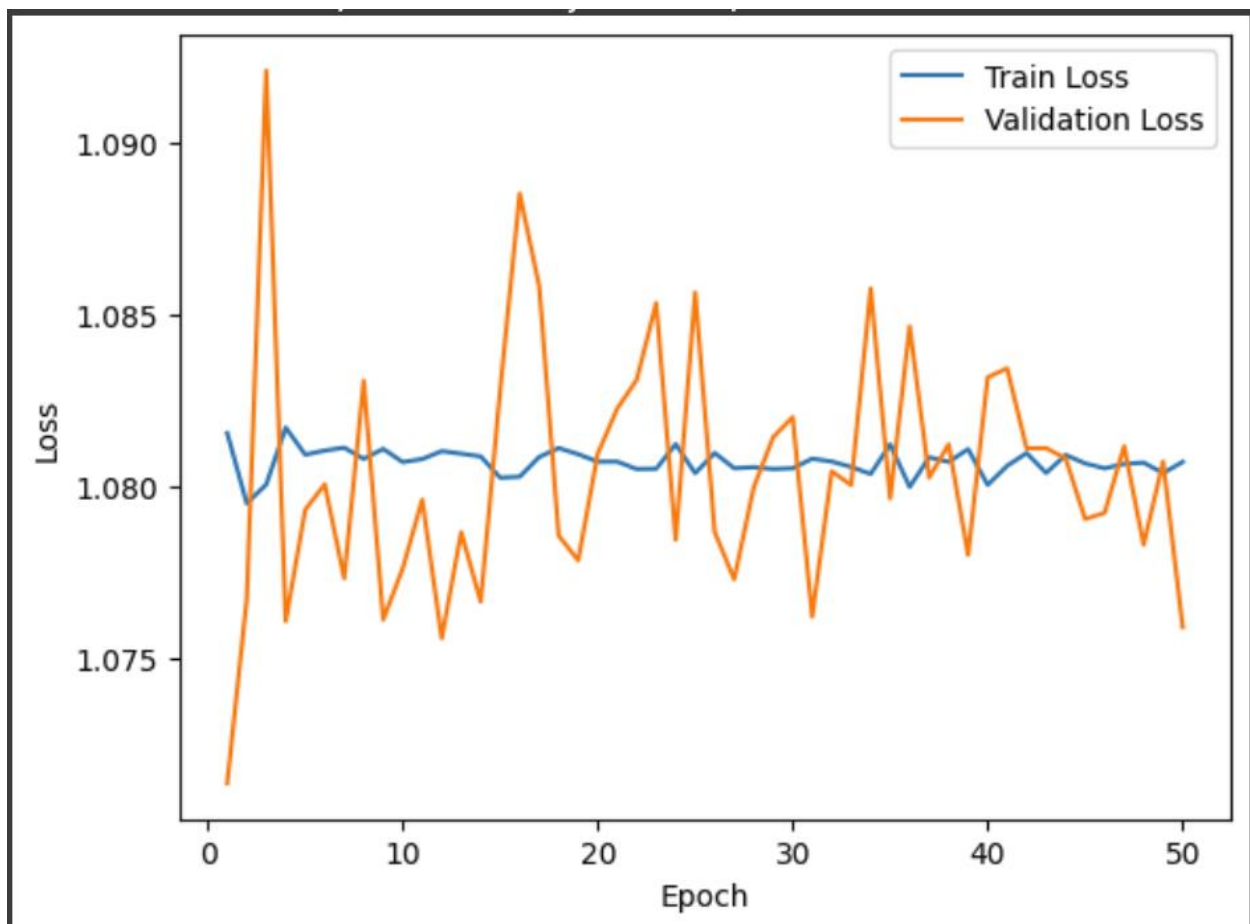
Epoch = 50

Normalization = True

Mini_batch_size_1

Model weight: Xavier

Learning Rate = 0.001

L2_lamba = 0.1

Optimizer: RMSProp
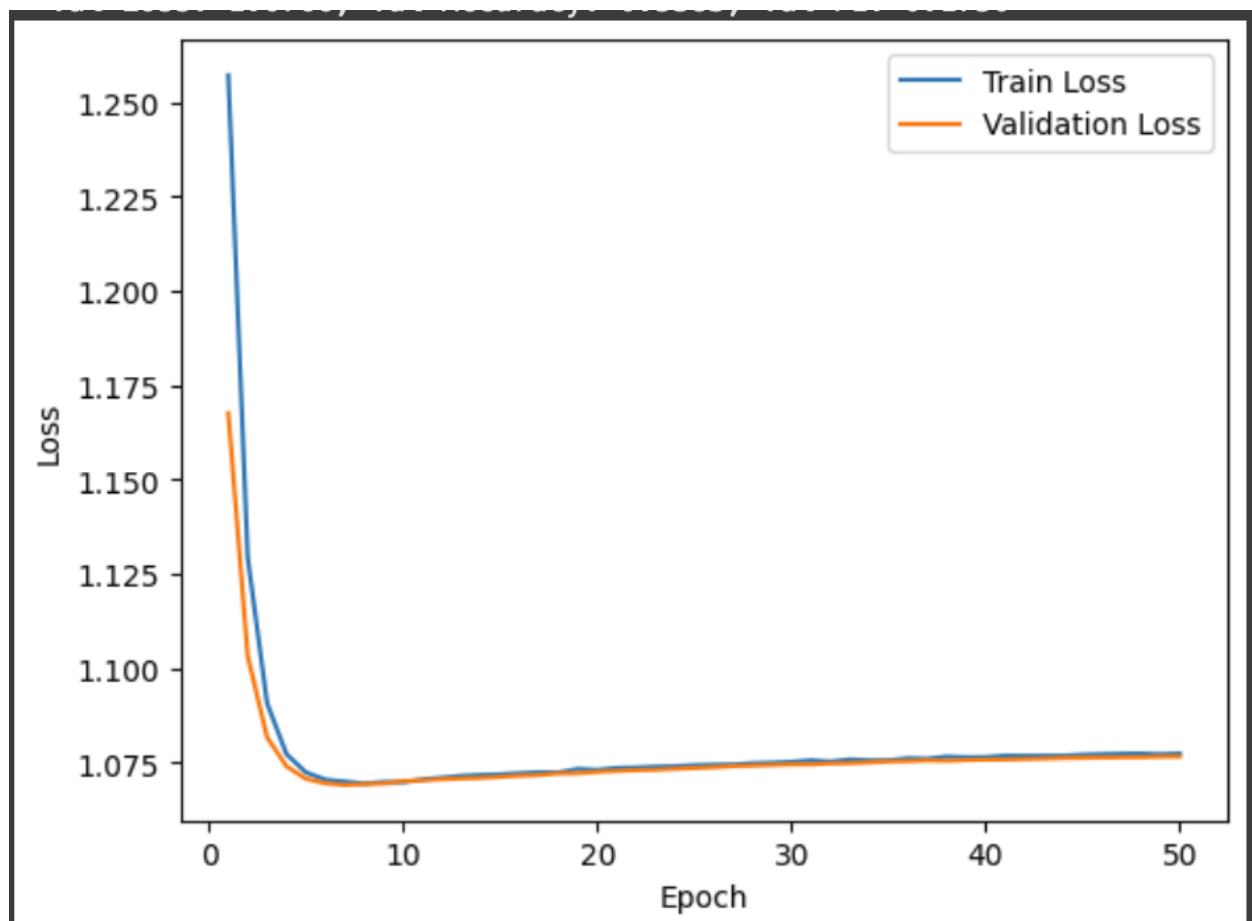
Epoch = 50

Normalization = True

Mini_batch_size_1

Model weight: Xavier

Learning Rate = 0.001

L2_lamba = 0.1

Optimizer: SGD

Epoch = 50

Normalization = True

Mini_batch_size_1

Model weight: Xavier

Learning Rate = 0.001

L2_lamba = 0.1

Optimizer: Momentum, Adam, RMSProp, SGD

Epoch = 50