FML : Assignment 02 Report



Department of Computer Science

Indian Institute of Technology Bombay

Guided by

**Prof. Sunita Sarawagi**

**Foundation of Machine Learning**

**CS 725**

Prepared By

**Bharat Patidar**

**23M0761**

**Rohit Singh Yadav**

**23M0773**

# Report :

In this Exercise/Assignment we have learnt classification on toy datasets with Feed Forward Neural Network using PyTorch Lightning. There are two datasets, one is a 4-class classification dataset simple and the other is digits dataset (10 classes) .

The optimal value of accuracy depends on the following three parameters.

1. Learning Rate

2. Epochs

3. Seed

The below table states the different parameter value for optimal accuracy :

|  |  |  |
| --- | --- | --- |
|  | Simple ( 4 classes ) | Digit ( 10 classes ) |
| Learning Rate | 0.1 | 0.01 |
| Epochs | 100 | 100 |
| Seed | 0 | 0 |
| Accuracy(%) | 99% | 98.25% |
| Loss | 0.027 | 0.094 |

# Report of Simple ( 4 Classes )

In the Feed Forward Neural Network Model the Dataset is taken from https://github.com/ashutoshbsathe/cs725-hw/tree/main/hw2 . In this model the Sample dataset is used where the number of classes were 4 .

We have implemented functions in classes : LitGenericClassifier & LitSimpleClassifier. LitSimpleClassifier is for the simple dataset and the LitGenericClassifier is for the both the models. We implemented the functions training\_step() , validation\_step() , test\_step() , predict() in LitGenericClassifier and transform\_input() , configure\_optimizer() in LitSimpleClassifier.

In this report we have to analyse the effect of the three parameters on the model’s accuracy and loss. As the parameters are :

1. Learning Rate

2. Epochs

3. Seed

**Effect of Learning Rate on Loss and Accuracy :**

As the Learning rate of the model decreases from the thresold value the step size or the loss increases . hence , the rate of change of learning rate directly affects the accuracy . Also the epochs is also a important parameter while finding the effeicient value of learning rate.

As in the calculation we have taken the epochs as 100 initially for calculating learning rate.

# Epochs = 100

Seed = 0

|  |  |  |
| --- | --- | --- |
| Learning Rate | Accuracy (%) | Loss |
| 1e-1 | 99% | 0.027 |
| 1e-2 | 99% | 0.041 |
| 1e-4 | 33% | 0.8 |
| 1e-6 | 33% | 1.34 |

**Conclusion :**

As the Accuracy is most optimal when we use Learning Rate = 1e-1 = 0.1 . As we decrease the size of the Learning Rate alpha the accuracy decreases . But with the decrease in learning rate the loss increases.

Hence , the most efficient Learning Rate is 0.1 as it gives the accuracy 99%.

**Effect of # of Epochs on Loss and Accuracy :**

As the Epochs of the model increases the number of iterations increases . hence , the rate of change of Epochs directly affects the accuracy . # of Epochs states that how much time the model will how much effort. Also the epochs is also a important parameter while finding the effeicient value of learning rate.

As in the calculation we have taken the Learning Rate as 0.1 for calculating # of Epochs.

Learning Rate = 0.1

Seed = 0

|  |  |  |
| --- | --- | --- |
| # of Epochs | Accuracy (%) | Loss |
| 100 | 99% | 0.027 |
| 250 | 99% | 0.04 |

**Conclusion :**

As the Accuracy is most optimal when we use Learning Rate = 1e-1 = 0.1 and use the # of Epochs = 100 . As we increase the number of the Epochs the accuracy remains same . But the increase in the Epochs causes the increase in loss.

Hence , the most efficient Learning Rate is 0.1 and Number of Epochs is 100.

# Report of Digit ( 10 Classes )

In the Feed Forward Neural Network Model the Dataset is taken from https://github.com/ashutoshbsathe/cs725-hw/tree/main/hw2 . In this model the Digit dataset is used where the number of classes were 10.

We have implemented functions in classes : LitGenericClassifier & LitDigitClassifier. LitDigitClassifier is for the Digit dataset and the LitGenericClassifier is for the both the models. We implemented the functions training\_step() , validation\_step() , test\_step() , predict() in LitGenericClassifier and transform\_input() , configure\_optimizer() in LitDigitClassifier.

In this report we have to analyse the effect of the three parameters on the model’s accuracy and loss. As the parameters are :

1. Learning Rate

2. Epochs

3. Seed

**Effect of Learning Rate on Loss and Accuracy :**

As the Learning rate of the model increases or decreases from the thresold value the step size or the loss increases and the accureacy decreases when we deviate from thresold value. hence , the rate of change of learning rate directly affects the accuracy . Also the epochs is also a important parameter while finding the effeicient value of learning rate.

As in the calculation we have taken the epochs as 100 initially for calculating learning rate.

# Epochs = 100

Seed = 0

|  |  |  |
| --- | --- | --- |
| Learning Rate | Accuracy (%) | Loss |
| 1e-1 | 10% | 2.3 |
| 1e-2 | 98.25% | 0.094 |
| 1e-4 | 95.99% | 1.25 |
| 1e-6 | 19% | 2.78 |

**Conclusion :**

As the Accuracy is most optimal when we use Learning Rate = 1e-2 = 0.01 . As we decrease the size of the Learning Rate alpha the accuracy decreases . Also when we increase the learning rate the accuracy drastically decreases also the loss increases as we increase or decrease the learning rate from 0.01.

Hence , the most efficient Learning Rate is **0.01**

**Effect of # of Epochs on Loss and Accuracy :**

As the Epochs of the model increases the number of iterations increases . hence , the rate of change of Epochs directly affects the accuracy . # of Epochs states that how much time the model will how much effort. Also the epochs is also a important parameter while finding the effeicient value of learning rate.

As in the calculation we have taken the Learning Rate as 0.01 for calculating # of Epochs.

Learning Rate = 0.01

Seed = 0

|  |  |  |
| --- | --- | --- |
| # of Epochs | Accuracy (%) | Loss |
| 100 | 98.25% | 0.094 |
| 250 | 98.25% | 0.097 |

**Conclusion :**

As the Accuracy is most optimal when we use Learning Rate = 1e-2 = 0.01 and use the # of Epochs = 100 . As we increase the number of the Epochs the accuracy remains same . But after a thresold value the accuracy remains constant. But the increase in the Epochs causes the increase in loss.

Hence , the most efficient Learning Rate is 0.01 and Number of Epochs is 100.

**a) The best epoch,**

The best value of # of epochs we found in both the dataset is 100.

after this value the accuracy is remaining constant but the loss increases as we increase # of epochs.

**b) Data-preprocessing performed (if any, write “not done” if you did not require any preprocessing)**

”Not Done” : we don’t find the need of data preprocessing.

**c) What did you do in order to prevent overfitting.**

For preventing Overfitting :

there are standard techniques :

1. Simplifying The Model

2. Early Stopping

3. Use Data Augmentation

4. Use Regularization

5. Use Dropouts.

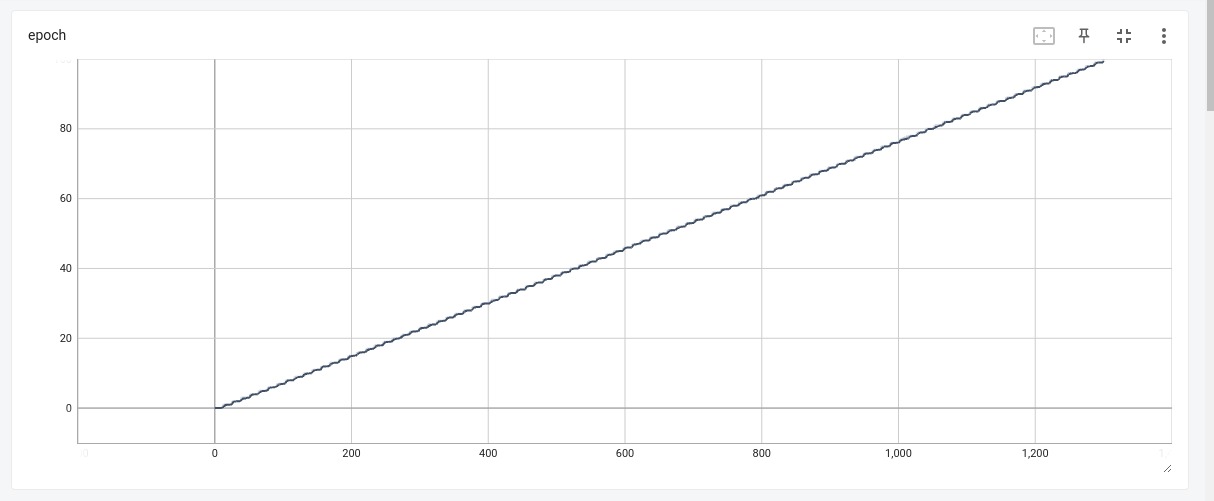
As we have used Early Stopping & Simplifying the model to prevent overfitting.

We have preffered adam model over the SGD(sophisticated gradient descent) .

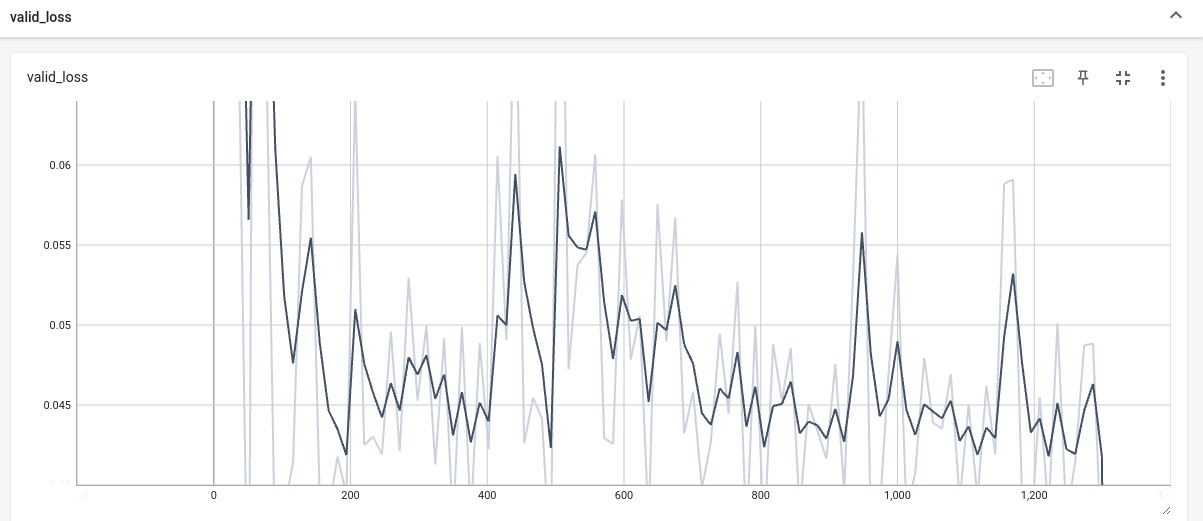
**GRAPHS :**

**Simple :**

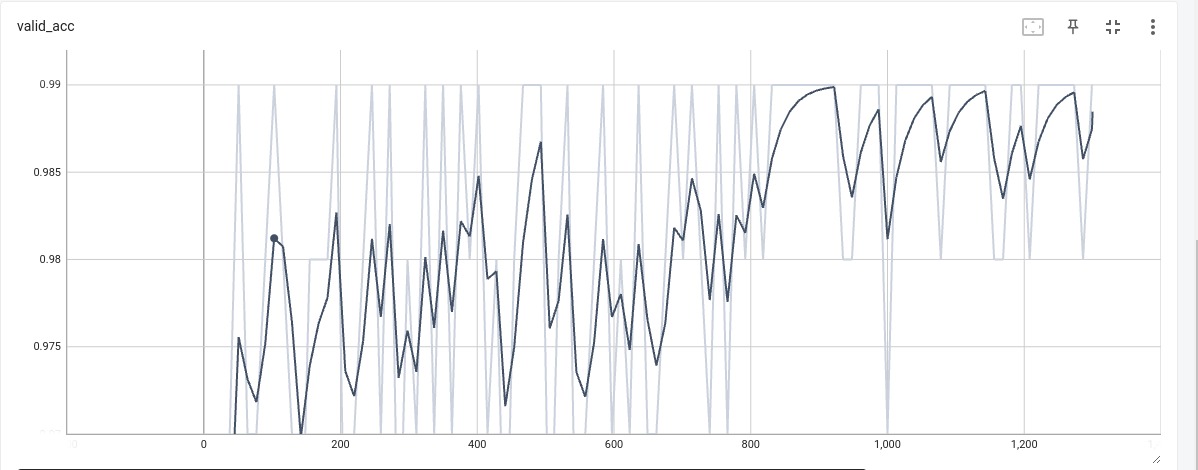
**1. Epochs**

****

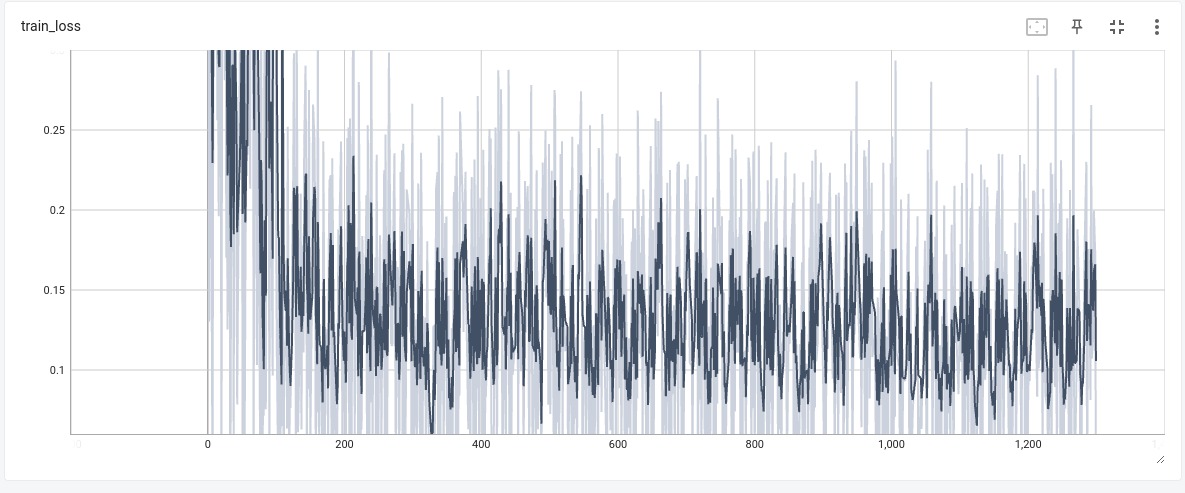
**2. Valid Loss**

****

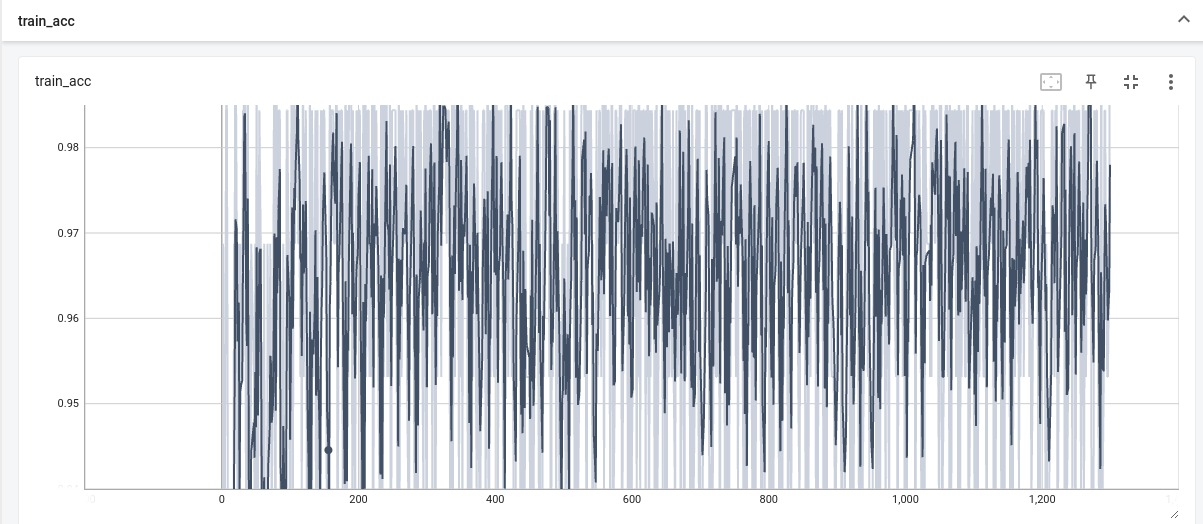
**3. Valid Accuracy**

****

**4. Train Loss**

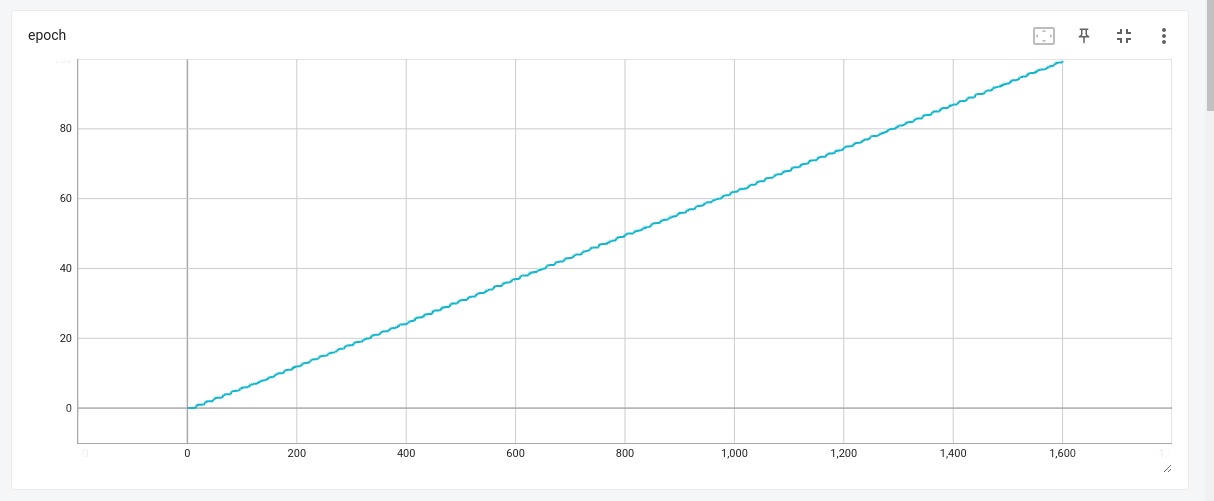
****

**5. Train Accuracy**

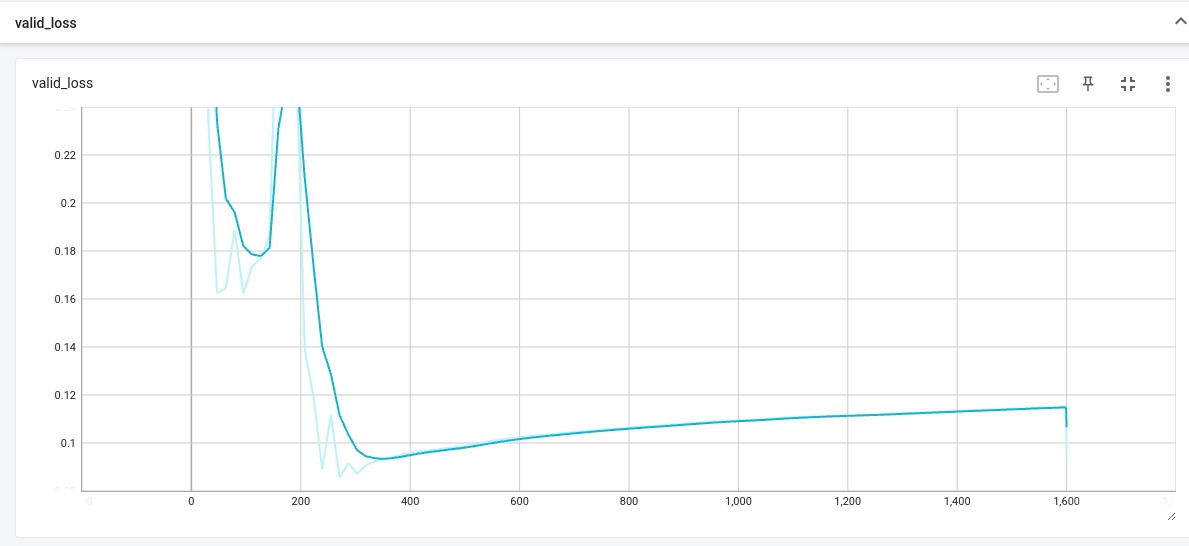


**Digit :**

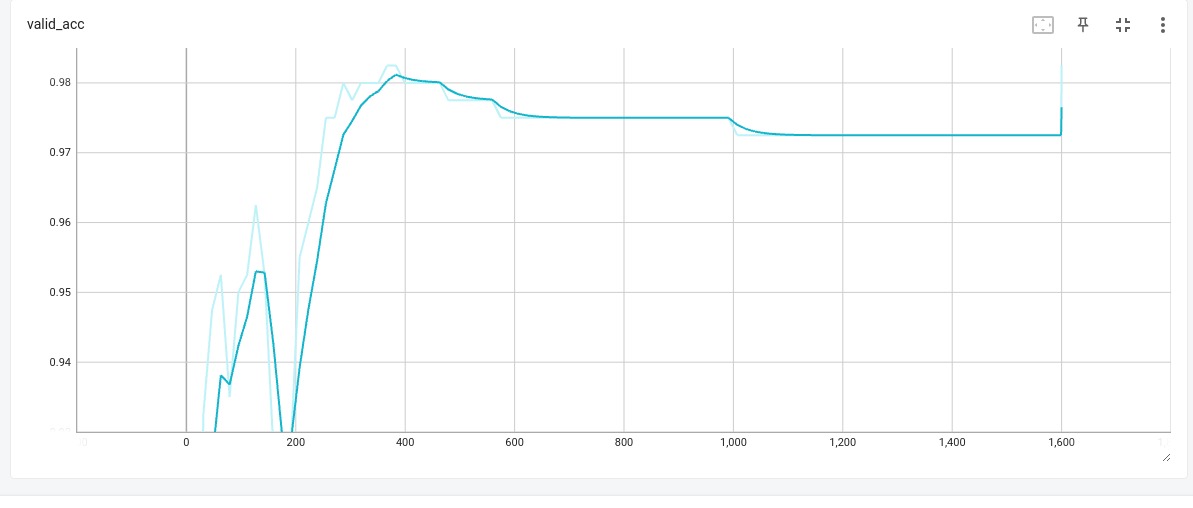
**1. Epochs**

****

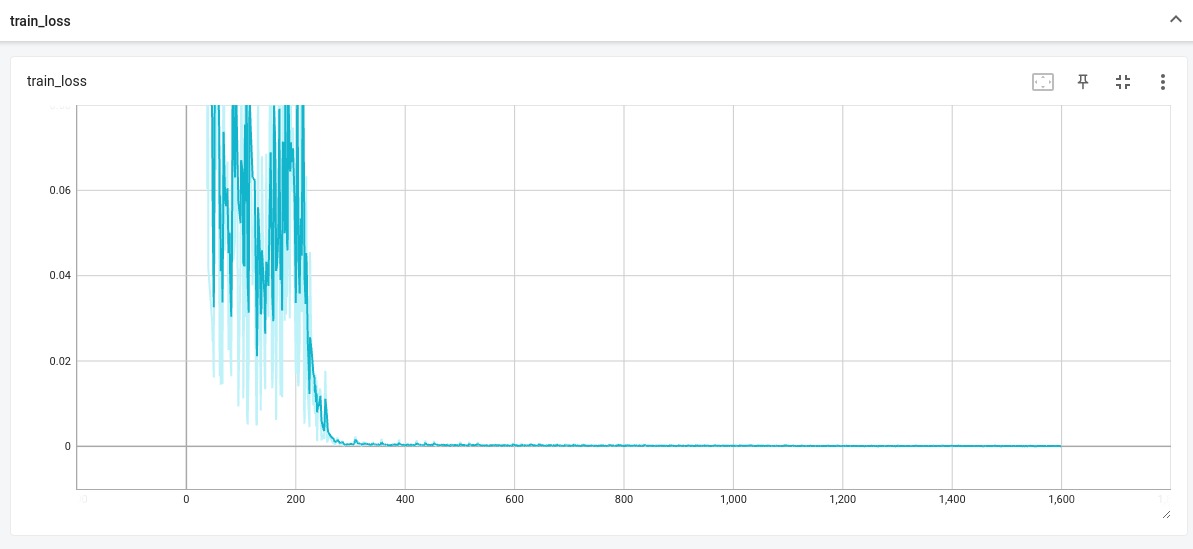
**2. Valid Loss**

****

**3. Valid Accuracy**

****

**4. Train Loss**

****

**5. Train Accuracy**

