

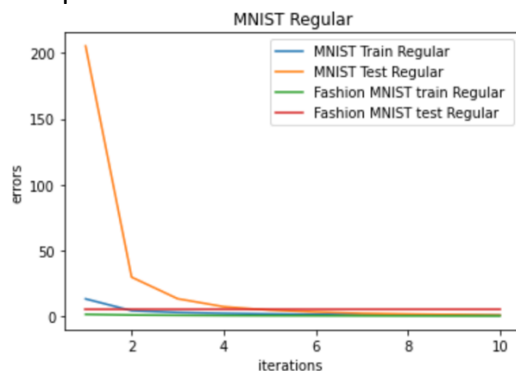
IST 597 Assignment 00010
Shikha Soneji
02/12/2022

Develop MNIST and Fashion MNIST model without Regularization

This model is created with SGD optimizer and the loss function is softmax cross entropy. This model is trained on GPU and the results for accuracy and Average cross entropy vary according to the datasets.

The graph below shows the metrics average Cross Entropy for training and testing data for MNIST and Fashion MNIST dataset for 10 trials.

It can be said that Fashion MNIST train seems to be consistent as well as the lowest metric as compared to other measures.



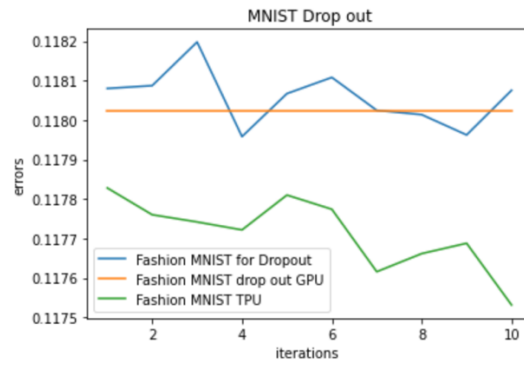
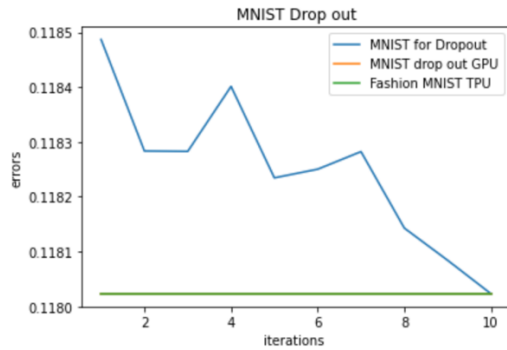
Inference for Test Data without Regularization

Test MSE: 3.0099
Accuracy: 10.0857
Standard Error: 4.5629
Variance: 83.2797
Test MSE: 3.0099

Develop MNIST and Fashion MNIST model with Dropout

This model is created with SGD optimizer and the loss function is softmax cross entropy. This model is trained on devices like GPU and TPU and the results for accuracy and Average cross entropy vary according to the devices. The output has dropout layers.

The graph below shows metrics for MNIST for average cross entropy with dropout. We see that the loss reduces for all the three.



Inference for Test Data with Dropout

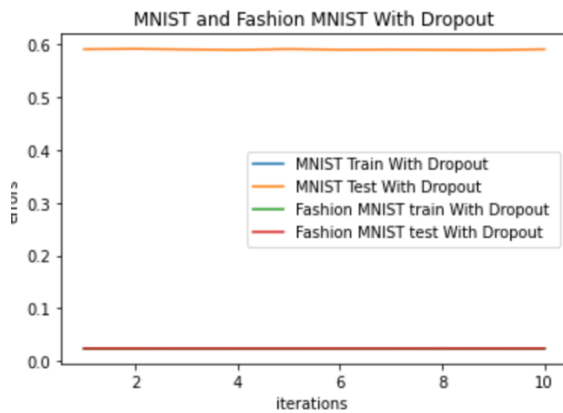
Test MSE: 0.0739

Accuracy: 9.7857

Standard Error: 0.1500

Variance: 0.0900

Test MSE: 0.0739

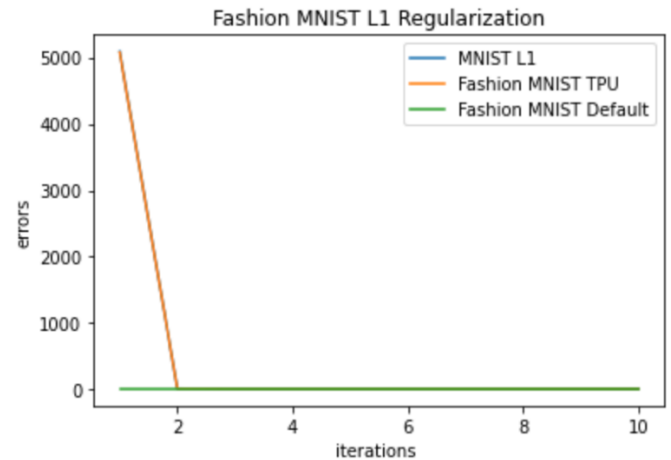
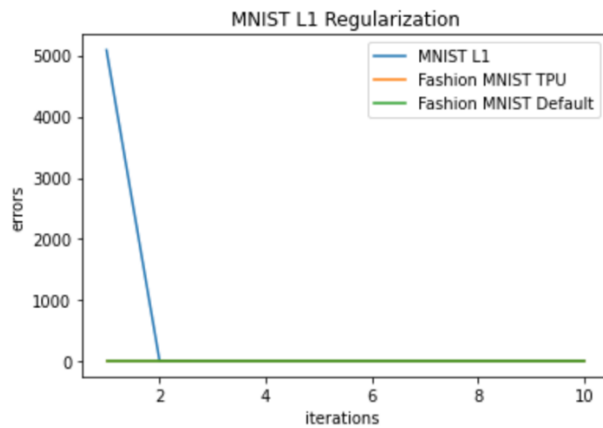


Develop MNIST and Fashion MNIST model with L1 Regularization

This model is created with SGD optimizer and the loss function is softmax cross entropy with L1 regularizer. This model is trained on devices like GPU and TPU and the results for accuracy and Average cross entropy vary according to the devices.

The graph below shows metrics for MNIST for average cross entropy with L1 Regularization.

We see that the loss reduces for all the three, however MNIST train data starts with a very high number and reduces in a steep fashion.



Inference for Test Data With L1 Regularization

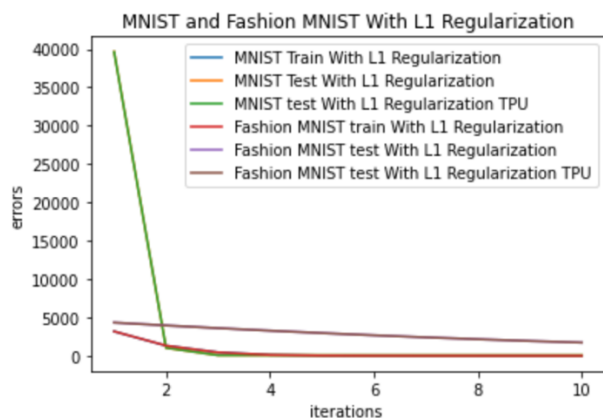
Test MSE: 5066.3594

Accuracy: 8.0571

Standard Error: 0.1500

Variance: 0.0900

Test MSE: 5066.3594

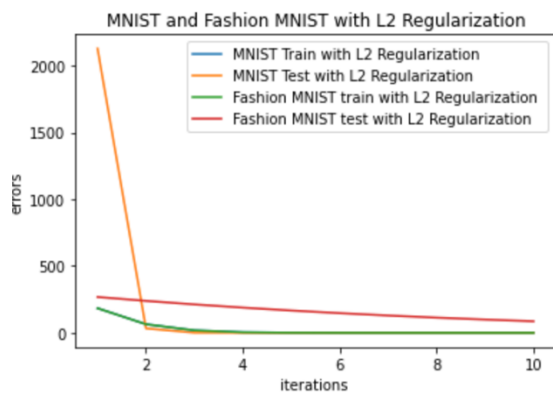
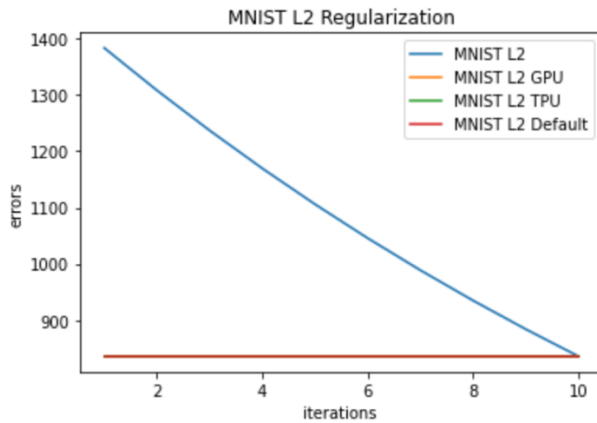


Develop MNIST and Fashion MNIST model with L2 Regularization

This model is created with SGD optimizer and the loss function is softmax cross entropy with L2 regularizer. This model is trained on devices like GPU and TPU and the results for accuracy and Average cross entropy vary according to the devices.

The graph below shows metrics for MNIST for average cross entropy with L2 Regularization.

We see that the loss reduces for all the three, however MNIST train data starts with a very high number and reduces in a steep fashion. For Fashion MNIST, CPU, TPU show a steep decrease.



Inference for Test Data with L2 Regularization
 Test MSE: 250.8078
 Accuracy: 8.9571
 Standard Error: 0.1500
 Variance: 0.0900
 Test MSE: 250.8078

Observations:

It is evident from the inference results that Dropout model works better as compared to other parameters. Also, TPU results are better than GPU results. The detailed accuracy and Average Cross Entropy for 10 trials are all listed in the colab notebooks in the folder provided.