
Assignment #10000

Amirreza Bagherzadeh
PennStateUniversity
PSUID: 946204639

Abstract

This assignment basically goes through the most basic RNN method, GRU, and a custom method MGU. Both these models are analyzed and the differences and the reason behind them is described.

1 RNN

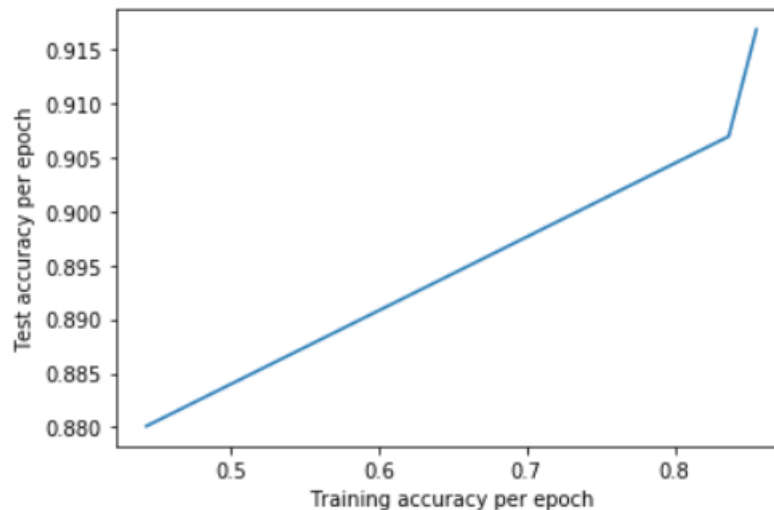
In assignment # 10000 we have 2 methods to analyze: GRU and MGU. Both has one hidden layer and the cells are based on first Basic LSTM file and they have been trained for 4 epochs. The dataset is Notmnist a dataset of english alphabet. We used large mnist for training and small mnist for test.

1.1 MGU

As the first experiment, we just run the model on MGU. The results aren't amazing but they are based on a hard dataset.

1.1.1 First time

Epoch	Training Accuracy	Test Accuracy	Loss
1	0.4428	0.8801	2.3556
2	0.8367	0.8977	2.1708
3	0.8588	0.9168	1.3110



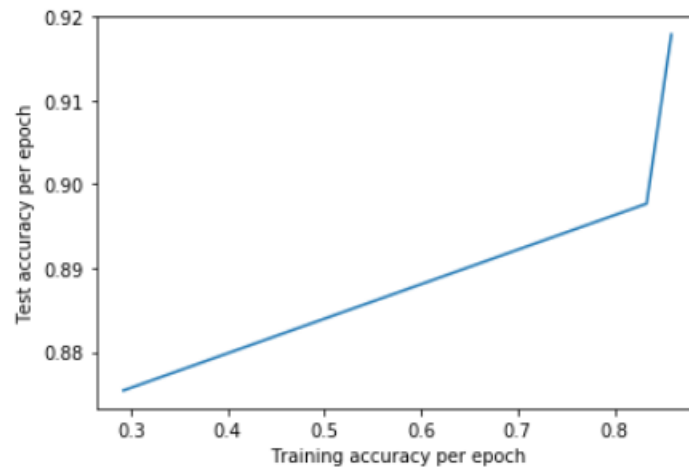
1.1.2 Second time

Epoch	Training Accuracy	Test Accuracy	Loss
1	0.4428	0.8801	2.3556
2	0.8367	0.8977	2.1708
3	0.8588	0.9168	1.3110

The result doesn't change from one test to another and as it is seen the results aren't that amazing though it seems like to be due to the training set being bias. Having the testing accuracy more than training empower this feeling.

1.1.3 GRU

Epoch	Training Accuracy	Test Accuracy	Loss
1	0.2924	0.8754	1.9049
2	0.8332	0.8977	0.5553
3	0.8588	0.9179	0.4618



As expected GRU has better performance than MGU. This was expected due to the fact that it has one less update cell and it has less flexibility when it comes to memorizing and forgetting. However it's way faster and more efficient for big datasets like this. This one doesn't also change by training it for multiple trials. To be honest it doesn't even make sense to train a model multiple time.