Assignment #00100 Report

Introduction

In this assignment, I will build a customized stochastic optimizer based on the example code provided by Ankur Mali [1]. The detailed code of the optimizer can be viewed at my github: https://github.com/Yuxuan-Liu-Eason/IST-597-00100

Then I will compare the customized optimizer with the keras inbuild optimizers (Adam, SGD, and RMSprop) on both datasets. There will be 12 models: 2 (with/without regularization) * 4 (customized, Adam, SGD, RMSprop) * 2 (MNIST and FMNIST).

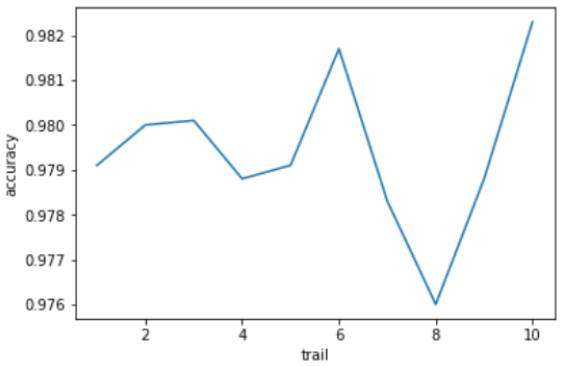
MNIST

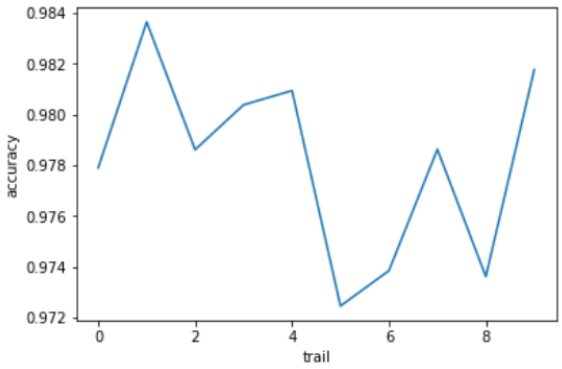
I used the best parameters from the previous assignment. The fixed parameters are: hidden_size1 = 300, hidden_size2 = 100, batch_size = 100, epoch = 20. Notice that different learning rates were applied to different models.

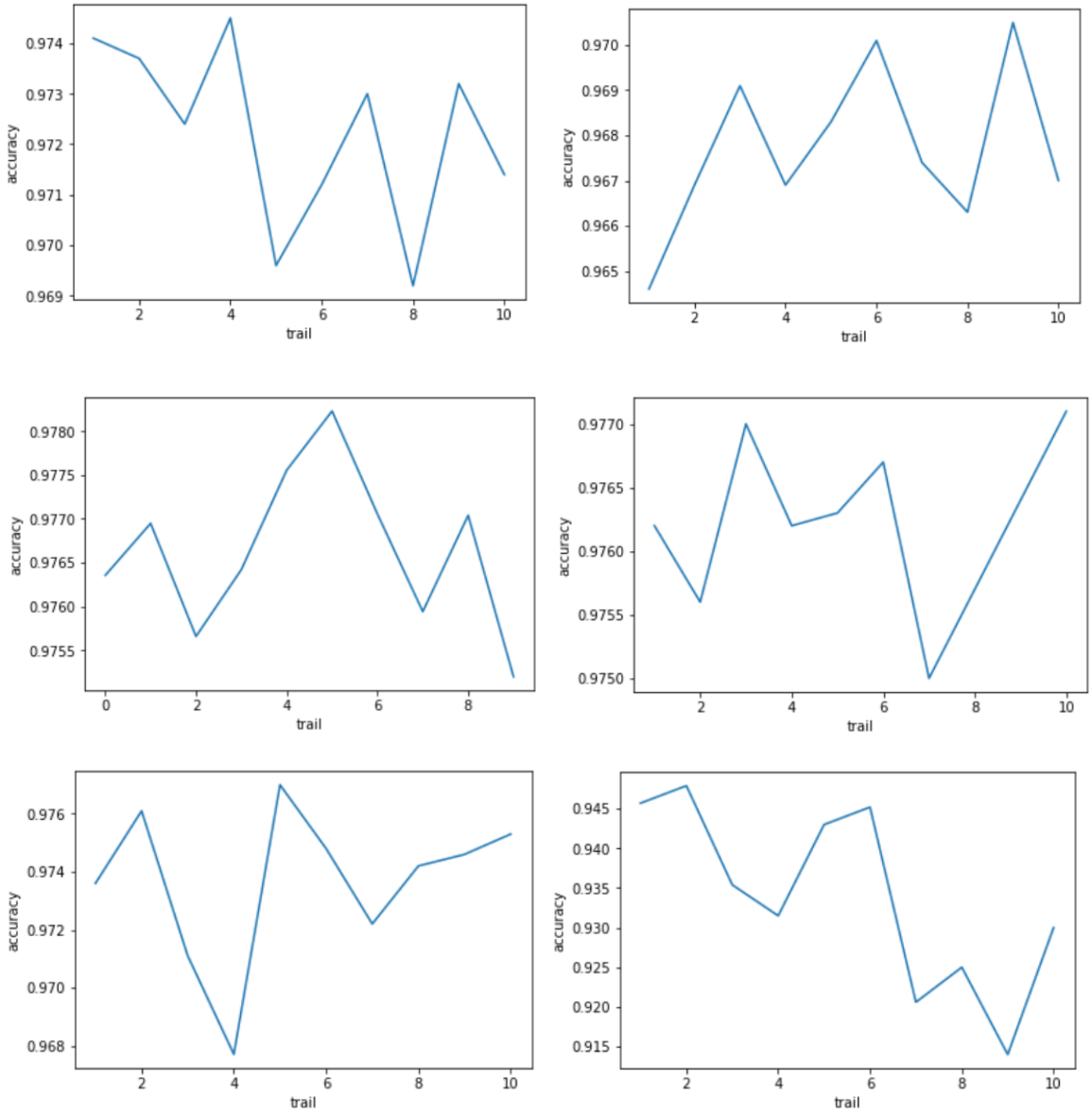
After running 10 trials for each model, I got the following results.

Learning rate	Regularization	Optimizer	Test accuracy	Standard error	Total time
0.001	None	Customized	0.9794	0.0017	3867
0.001	L2	Customized	0.977	0.0036	4406
0.005	None	Adam	0.9749	0.0013	3883
0.001	L2	Adam	0.9677	0.0017	4383
0.1	None	SGD	0.9763	0.0007	3002
0.1	L2	SGD	0.9762	0.0006	3297
0.001	None	RMSprop	0.9737	0.0026	3943
0.001	L2	RMSprop	0.9338	0.0110	4384

The following plots show their variabilities during the 10 trials.







Robustness: We can see that the customized optimizer without regularization gives the best accuracy. Actually, all of them have relatively good accuracy except for RMSprop with l2 regularization. The l2 in RMSprop decreased the accuracy dramatically. I am not sure why this happened. The l2 penalty in Adam also decreased the accuracy by a little. Overall, the customized optimizer has the best accuracy, then is SGD, Adam and finally RMSprop.

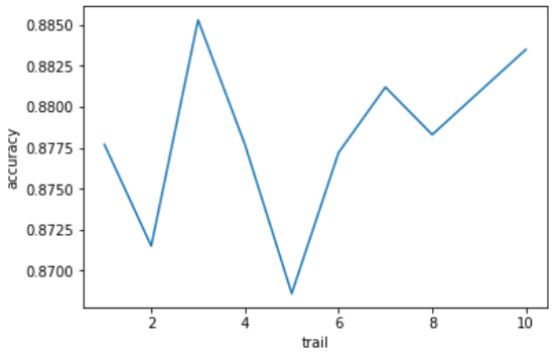
Stability: SGD has the best stability because the standard error during 10 trials is the smallest. Adam is the second stable. Then is the customized. The l2 penalty leads to more instability. The least stable is RMSprop. The l2 penalty also decreases the stability.

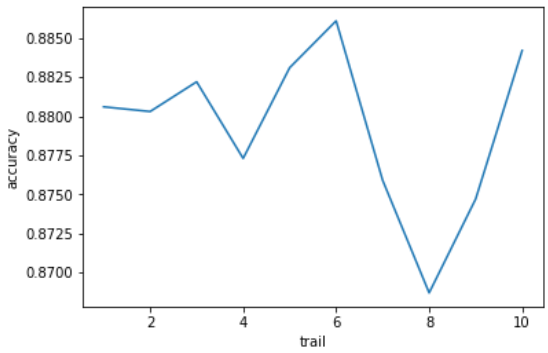
Speed: According to total running time, SGD is the fastest. The remaining 3 optimizers have similar speed. Also, l2 penalty always increase the calculation time.

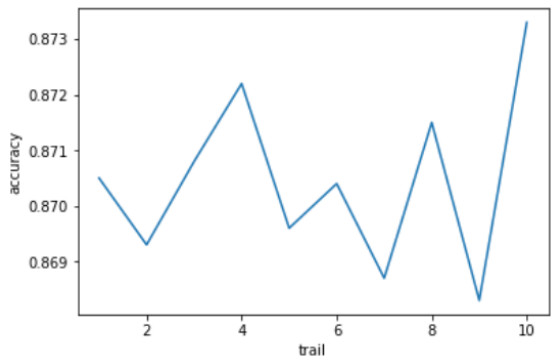
Fashion MNIST

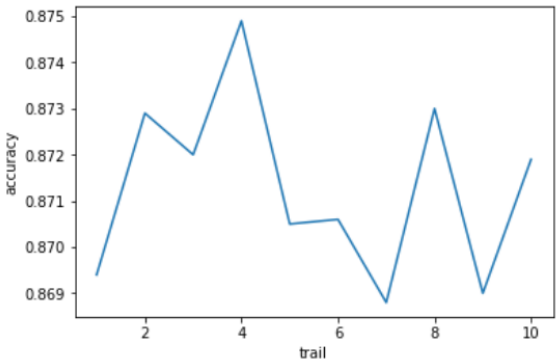
Same parameters were applied to the fashion mnist dataset. After running 10 trials for each model, I got the following results.

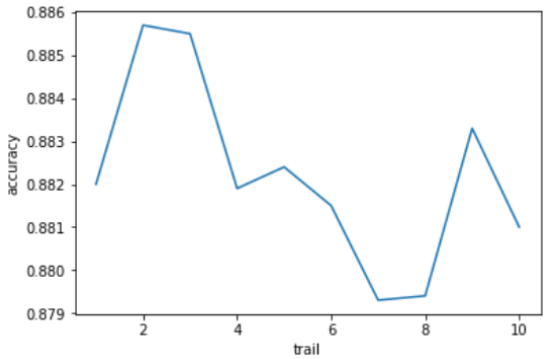
Learning rate	Regularization	Optimizer	Accuracy	Standard error	Total time
0.001	None	Customized	0.8782	0.0048	3969
0.001	L2	Customized	0.8793	0.0049	4312
0.0001	None	Adam	0.8705	0.0015	3704
0.0001	L2	Adam	0.8713	0.0019	4058
0.1	None	SGD	0.8822	0.0021	2987
0.1	L2	SGD	0.8829	0.0017	3260
0.001	None	RMSprop	0.8025	0.0285	3948
0.001	L2	RMSprop	0.8125	0.0082	4240

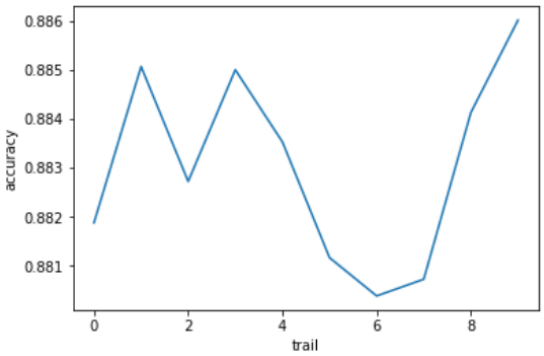


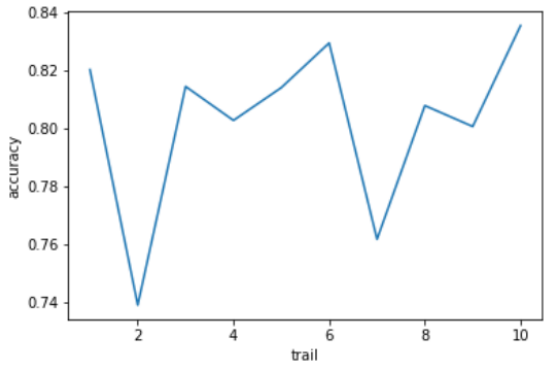


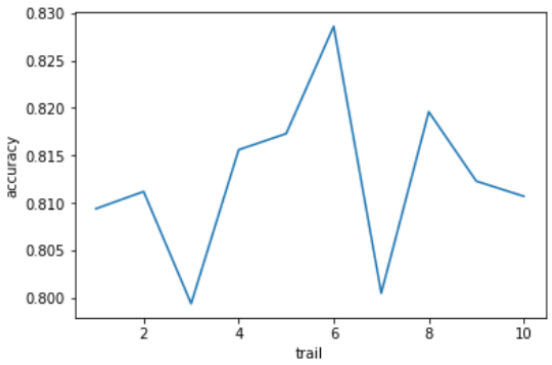












Robustness: We can see that SGD gives the best accuracy. The next is the customized optimizer. The l2 penalty helps increase the accuracy a little bit. Adam without regularization gives relatively worse performance. With l2 penalty, the accuracy increases by a little. Among the three optimizers, RMSprop has the worse performance.

Stability: Both Adam and SGD have pretty good stability because the standard error during 10 trials is small. The customized optimizer has relatively worse stability. RMSprop has the worst stability.

Speed: SGD is the fastest. The customized optimizer and RMSprop are the slowest. Adam is in between.

Summary

To compare the three optimizers, SGD has both relatively good accuracy and best speed. The customized optimizer also did a very good job. However, this statement is not necessarily true. According to what I learned, the Adam should give better accuracy than SGD. So it is possible that the models with Adam and RMSprop are not at their optimal performance.

Reference

[1]
https://github.com/AnkurMali/IST597_Spring_2022/blob/main/assignment00100/Spring2022_IST597_week5_collab_custom_optimizer.ipynb