Assignment 1 Writeup

- Name: Jieun Seong
- GT Email: jseong8@gatech.edu
- GT ID: 903004701

Two-Layer Neural Network

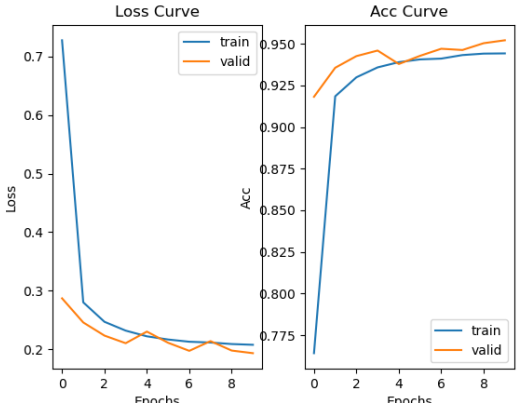
1. Learning Rates

Tune the learning rate of the model with all other default hyper-parameters fixed. Fill in the table below:

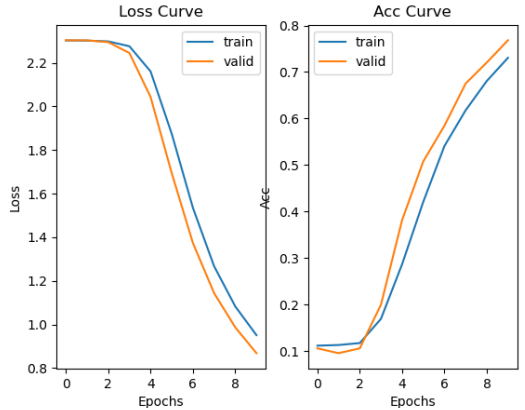
	lr=1	lr=1e-1	lr=1e-2	lr=5e-2
Training Accuracy	0.9442	0.9214	0.7306	0.9088
Test Accuracy	0.9482	0.9267	0.7573	0.9150

1. Learning Curve

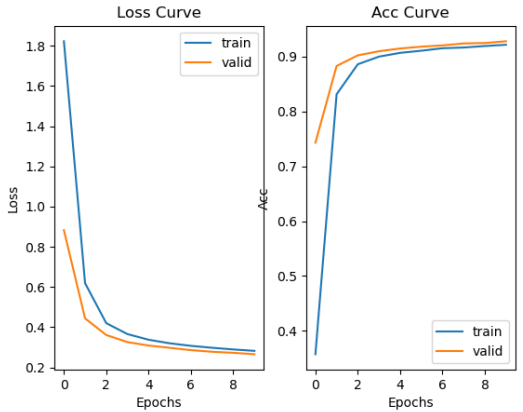
Plot the learning curves [both ac



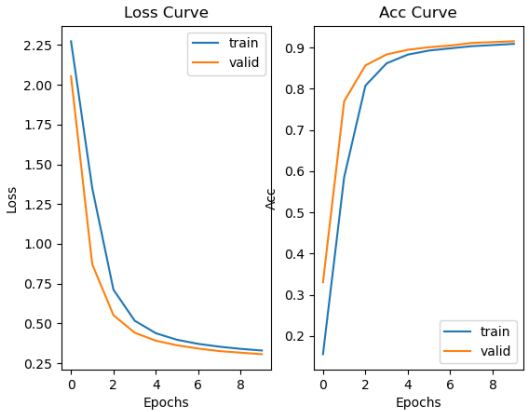
learning_rate = 1



learning_rate = 1e-2



learning_rate = 1e-1



learning_rate = 5e-2

1. Learning Rates

Describe and Explain your findings:

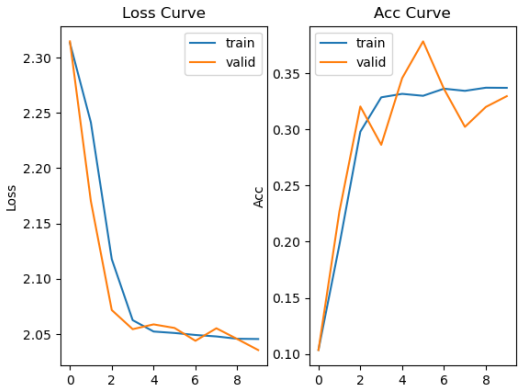
It’s difficult to say that the rate of improvement is positively or negatively related to the learning rates. Even though the algorithm will reach some optima, it’s not guaranteed that the optima will be the global optima. Therefore, there can unfortunately be some values for learning rates that can lead to non-global optima and get stuck there. lr=1e-2 is such a case in our experiment.

2. Regularization

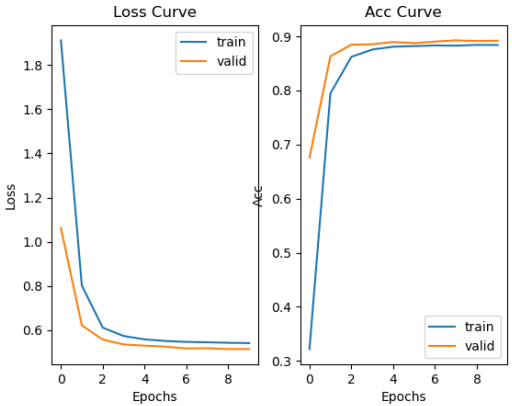
Tune the regularization coefficient of the model with all other default hyper-parameters fixed. Fill in the table below:

	alpha=1e-1	alpha=1e-2	alpha=1e-3	alpha=1e-4	alpha=1e-0
Training Accuracy	0.3371	0.8841	0.9214	0.9294	0.1044
Validation Accuracy	0.3297	0.8920	0.9277	0.9338	0.1060
Test Accuracy	0.3756	0.8920	0.9267	0.9334	0.1028

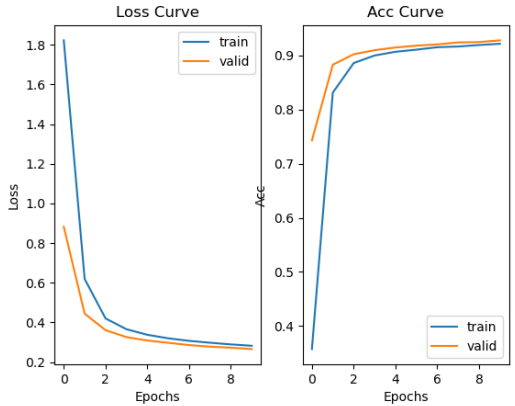
2. Regularization



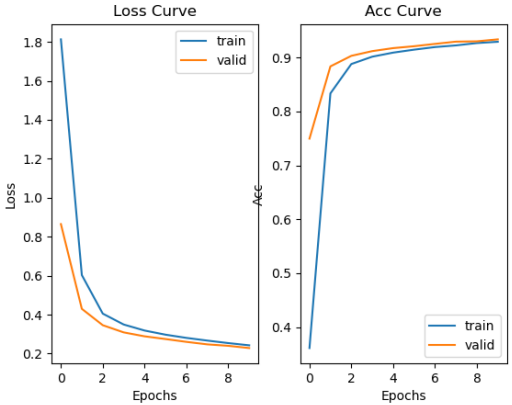
Alpha = 1e-1



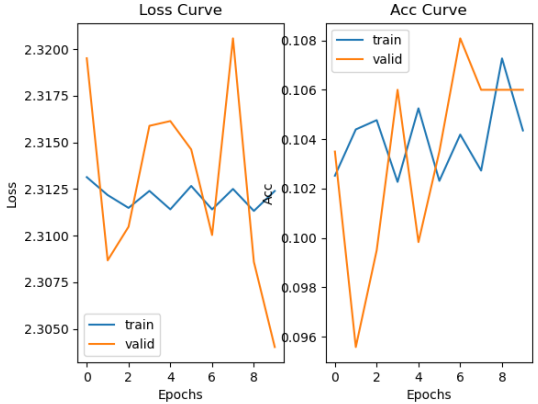
Alpha = 1e-2



Alpha = 1e-3



Alpha = 1e-4



Alpha = 1e-0

2. Regularization

Describe and Explain your findings:

Based on observation, the bigger the regularization constant was, the worse the accuracy got.

Regularizations are methods to reduce the error by appropriately fitting parameters and avoid overfitting. As we observed in the data and the plots, when we increased the regularization constant, the accuracies fluctuated a lot, giving too much flexibility to the model.

3. Hyper-parameter Tuning

alpha	lr	Training accuracy	Validation Accuracy	Test Accuracy
1E-04	1	0.9788	0.9712	0.9696
1E-04	2	0.9691	0.9664	0.9666
1E-04	1.5	0.9752	0.9669	0.9679
1E-04	0.5	0.9721	0.9677	0.9668
1E-05	1	0.9833	0.9717	0.9713
5E-05	1	0.9813	0.9712	0.9707
1E-06	**1**	**0.9838**	**0.9712**	**0.9713**
1E-07	1	0.9838	0.9712	0.9713

meters for better accuracy. Create a table below and
and accuracy into the table:

Briefly explain why your choice works:

From previous observations, we saw that alpha=1e-4 and lr=1 work well, while larger values for alpha and there were higher chances that smaller values of lr would not work as well. Thus, I tried the values for alpha and lr around alpha=1e-4 and lr=1. The accuracies more or less had a plateau at around alpha=1e-6 and lr=1.