

# MACHINE LEARNING GLASSES RESULTS LOG

KIRK SWANSON

1. 11/26/2017

- (1) Merged recent updates to GitHub
- (2) Downloaded images from a while ago. This first one is batch size 80, learning rate 0.001, beta 0.01, dropout 0.5, including the recent change that each batch is chosen from a newly randomly chuffed metadata:

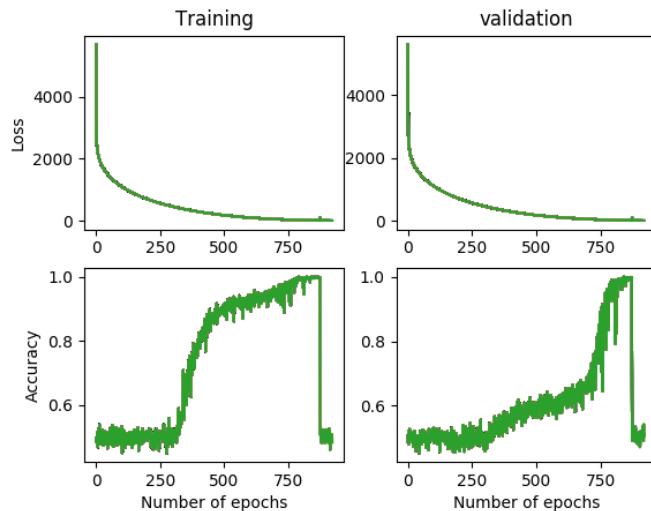


FIGURE 1. Batch Size = 80, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.5

- (3) This run had an optimal validation accuracy of 99.7635 and an optimal test accuracy of 99.7297. I also tested this multiple times, running the SAME set of hyperparameters, to see what is going on with that weird spike. Does it keep showing up? What is going on? Here are more examples, all run at the same hyperparameters as above:

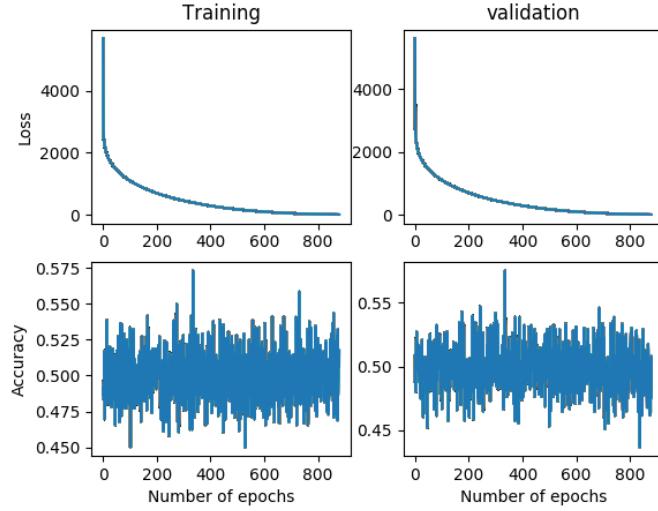


FIGURE 2. Batch Size = 80, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.5

- (4) This run had an optimal validation accuracy of 47.0608, optimal test accuracy of 47.6689, and final validation accuracy of 49.223, test accuracy of 50.0338

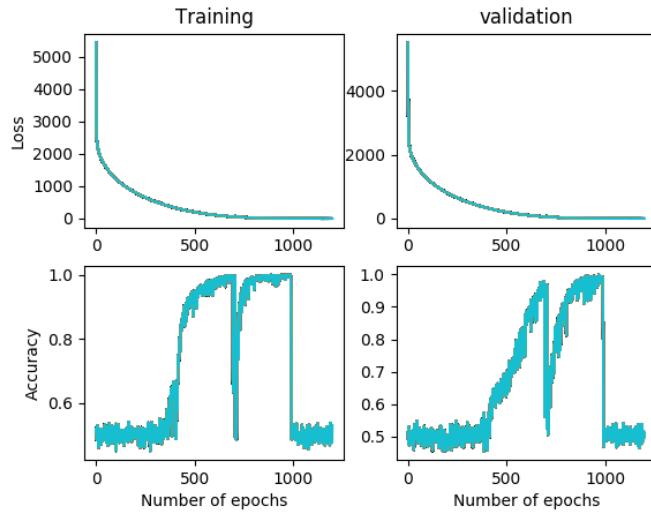


FIGURE 3. Batch Size = 80, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.5

- (5) This run had an optimal validation accuracy of 99.4595, optimal test accuracy of 99.3243, and final validation accuracy of 50.777, test accuracy of 49.9662.

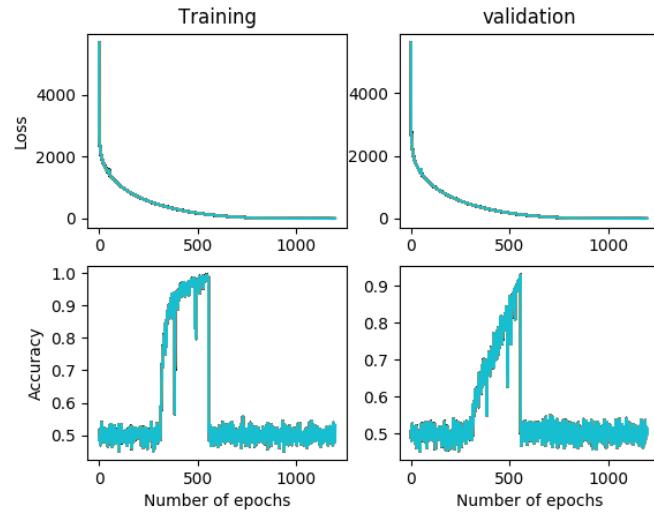


FIGURE 4. Batch Size = 80, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.5

- (6) This run had an optimal validation accuracy of 992.5, optimal test accuracy of 91.7568, and final validation accuracy of 49.223, test accuracy of 50.0338.

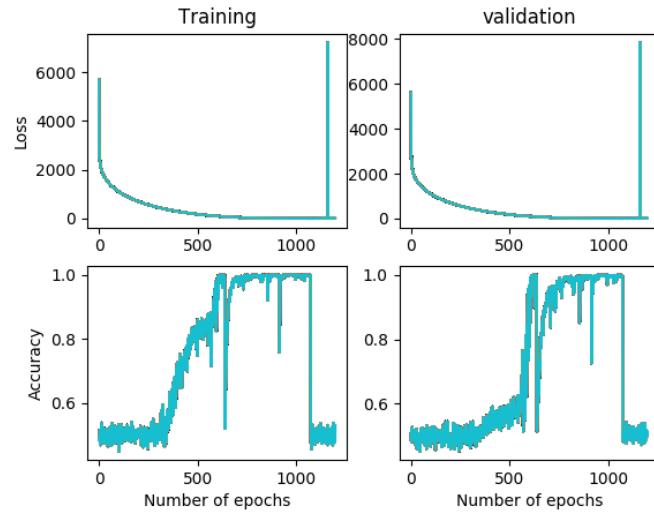


FIGURE 5. Batch Size = 80, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.5

- (7) This run had an optimal validation accuracy of 99.8987, optimal test accuracy of 99.8649, and final validation accuracy of 50.777, test accuracy of 49.9662.

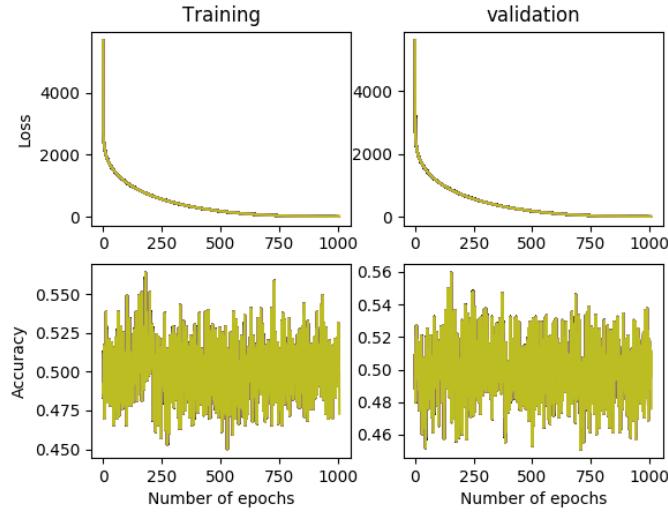


FIGURE 6. Batch Size = 80, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.5

- (8) This run had an optimal validation accuracy of 55.4054, optimal test accuracy of 54.223, and final validation accuracy of 49,223, test accuracy of 50.0338.
- (9) The following are images of runs on the dataset that consists of glasses and liquids that are each 12 time steps away from the assumed glass transition temperature of 0.21.

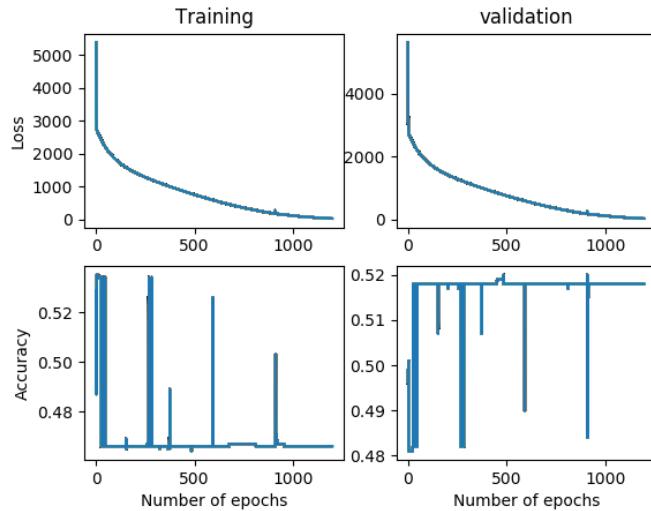


FIGURE 7. Batch Size = 100, Learning Rate = 1e-4, Beta = 0.01, Dropout = 0.5

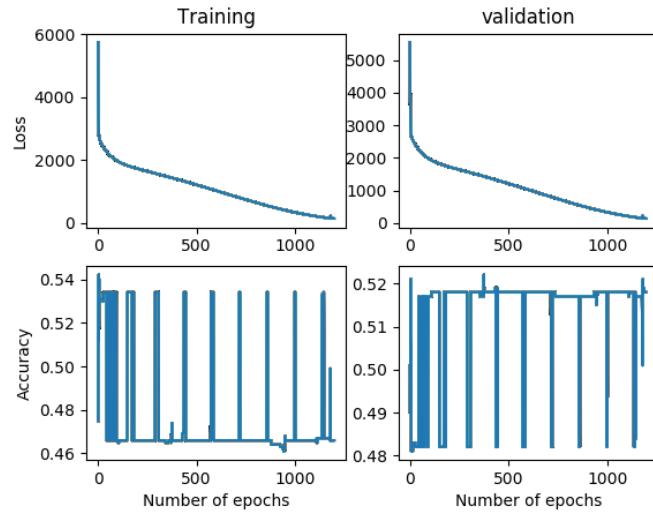


FIGURE 8. Batch Size = 10, Learning Rate = 1e-4, Beta = 0.01, Dropout = 0.5

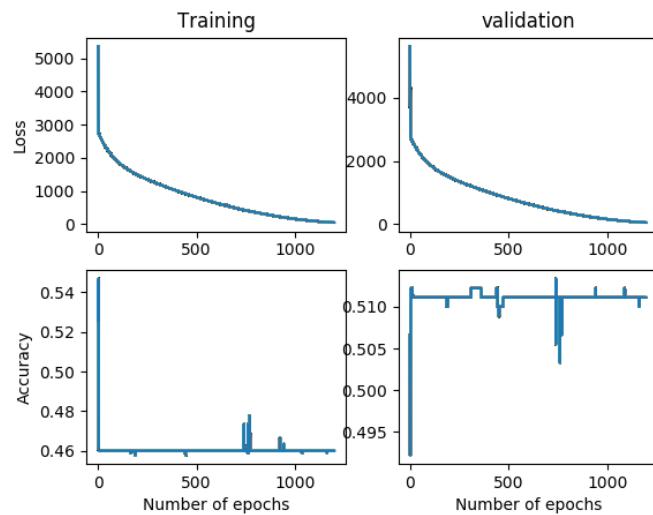


FIGURE 9. Batch Size = 150, Learning Rate = 1e-4, Beta = 0.01, Dropout = 0.5

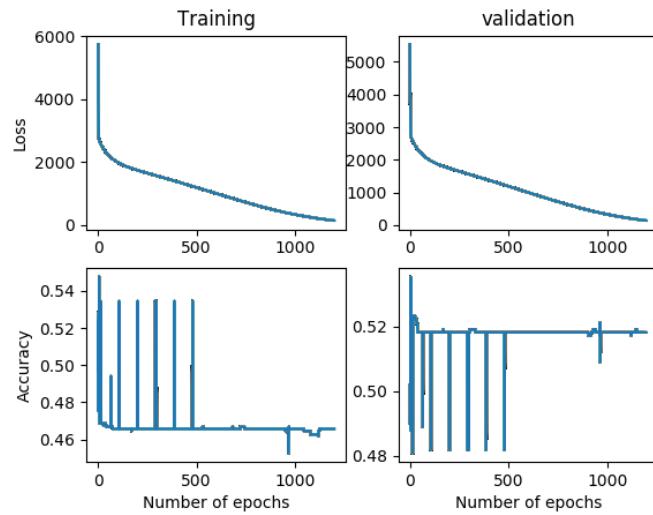


FIGURE 10. Batch Size = 15, Learning Rate = 1e-4, Beta = 0.01, Dropout = 0.5

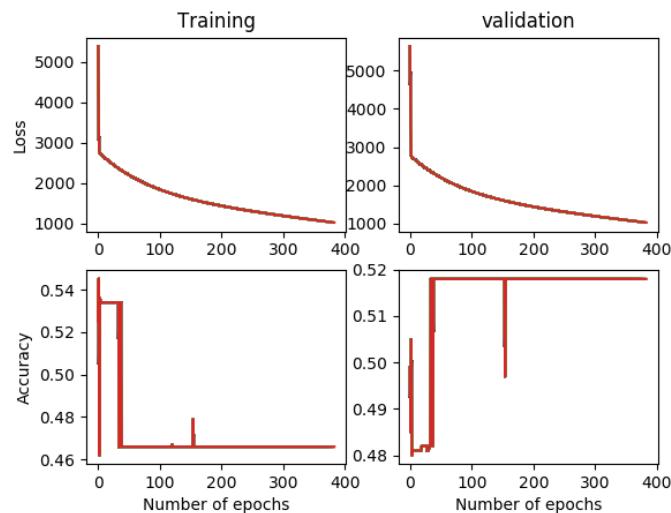


FIGURE 11. Batch Size = 200, Learning Rate = 1e-4, Beta = 0.01, Dropout = 0.5

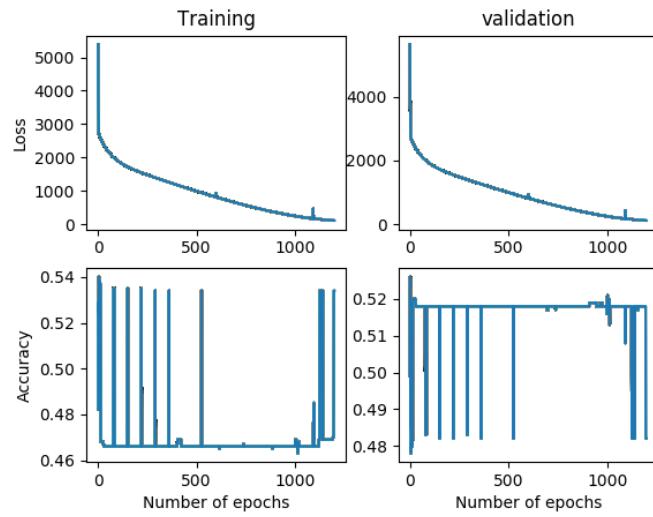


FIGURE 12. Batch Size = 20, Learning Rate = 1e-4, Beta = 0.01, Dropout = 0.5

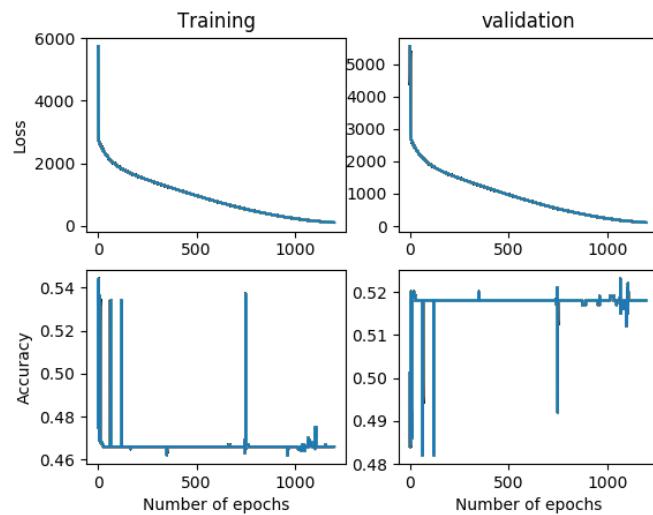


FIGURE 13. Batch Size = 25, Learning Rate = 1e-4, Beta = 0.01, Dropout = 0.5

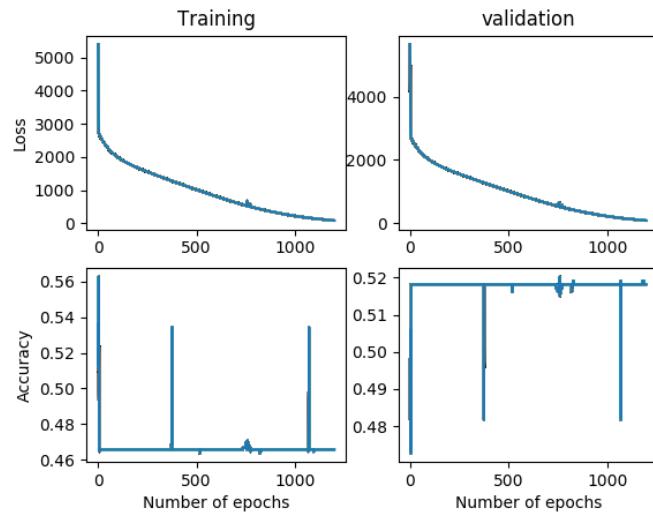


FIGURE 14. Batch Size = 30, Learning Rate =  $1e-4$ , Beta = 0.01, Dropout = 0.5

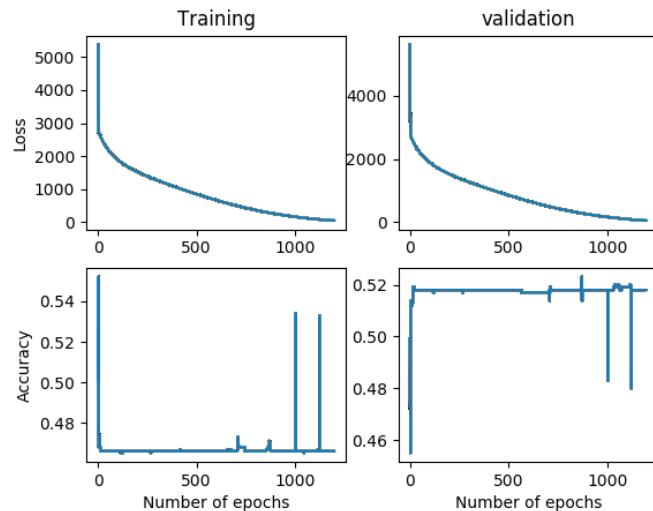


FIGURE 15. Batch Size = 50, Learning Rate =  $1e-4$ , Beta = 0.01, Dropout = 0.5

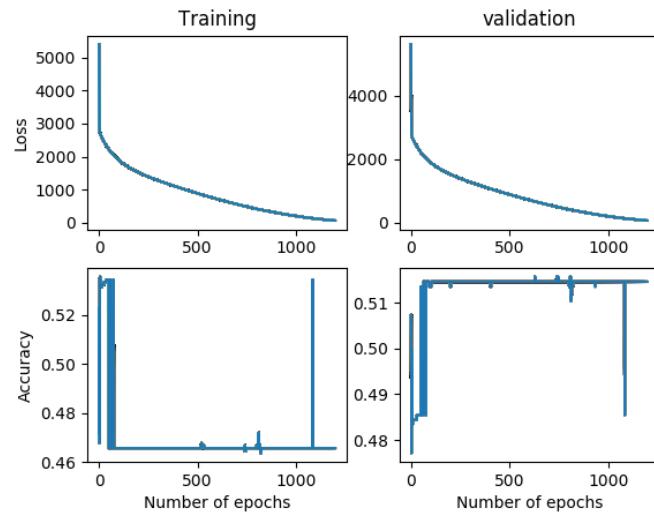


FIGURE 16. Batch Size = 80, Learning Rate = 1e-4, Beta = 0.01, Dropout = 0.5

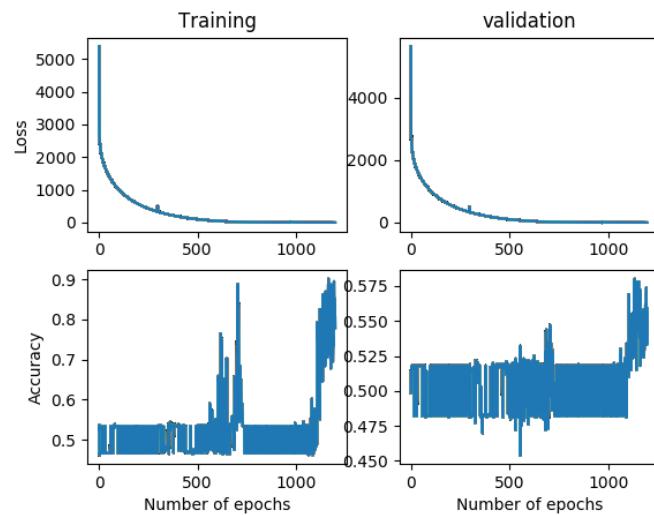


FIGURE 17. Batch Size = 100, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.5

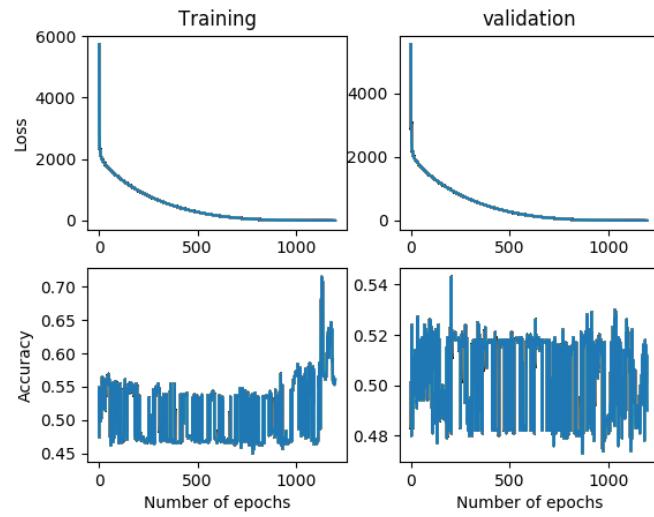


FIGURE 18. Batch Size = 10, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.5

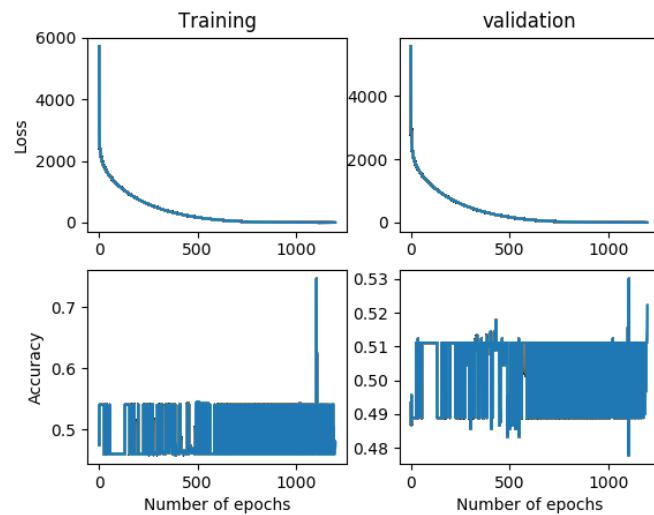


FIGURE 19. Batch Size = 150, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.5

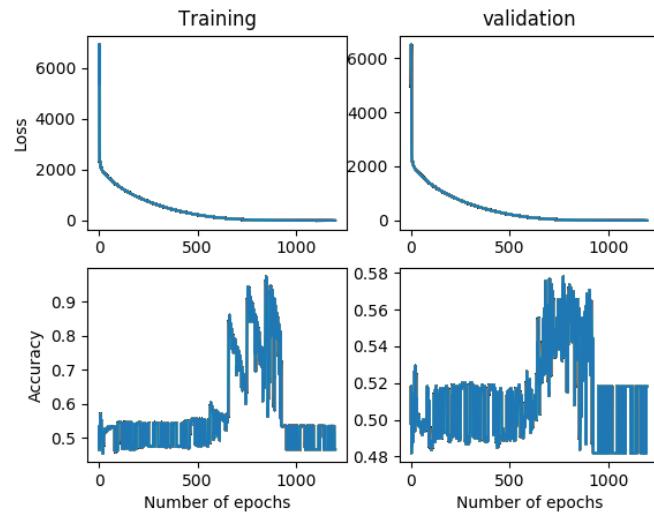


FIGURE 20. Batch Size = 15, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.5

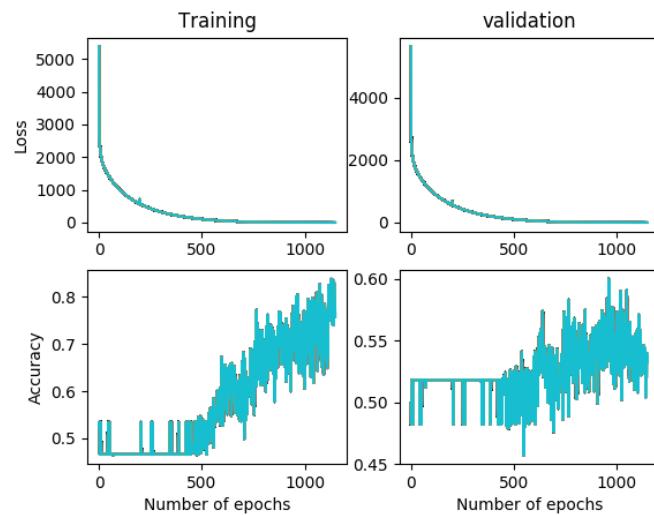


FIGURE 21. Batch Size = 200, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.1

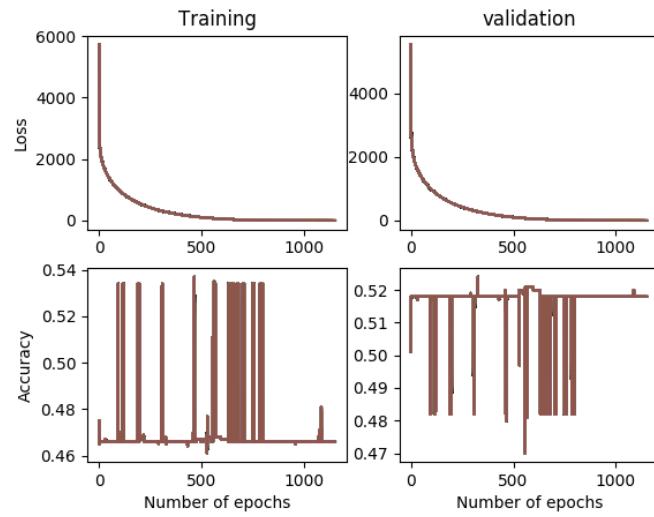


FIGURE 22. Batch Size = 200, Learning Rate = 1e-3, Beta = 0.01,  
Dropout = 0.2

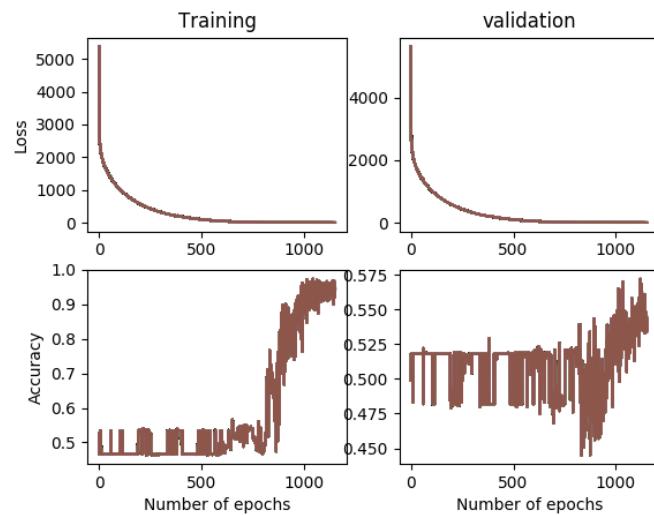


FIGURE 23. Batch Size = 200, Learning Rate = 1e-3, Beta = 0.01,  
Dropout = 0.3

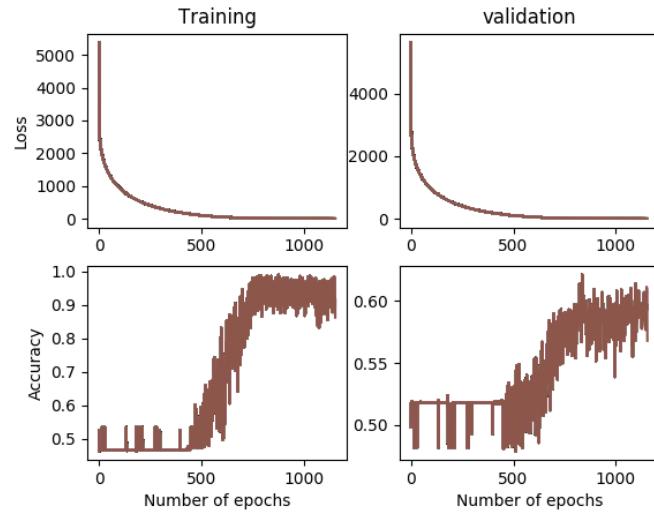


FIGURE 24. Batch Size = 200, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.4

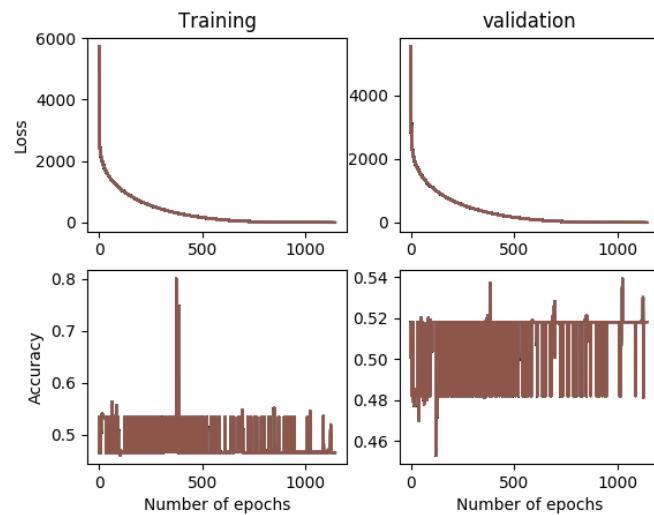


FIGURE 25. Batch Size = 200, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.5

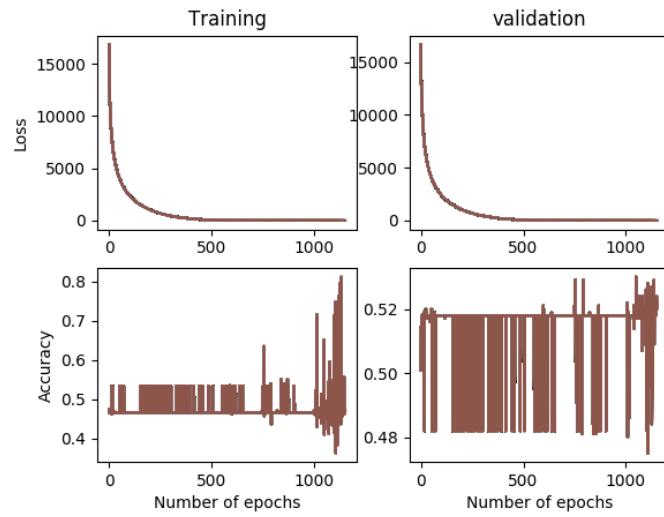


FIGURE 26. Batch Size = 200, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.3

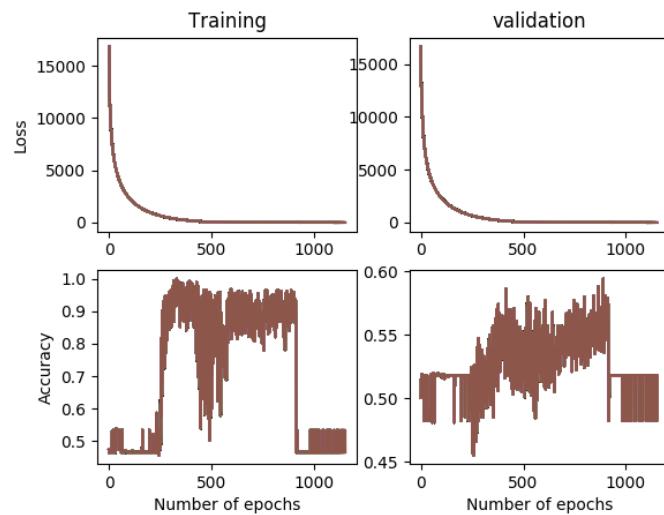


FIGURE 27. Batch Size = 200, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.4

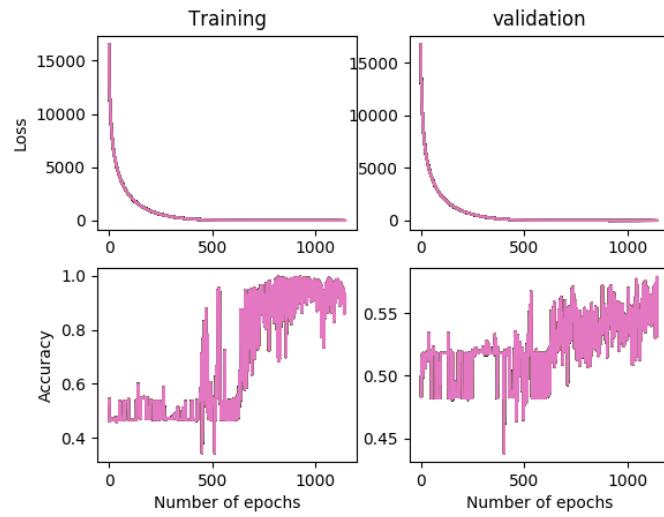


FIGURE 28. Batch Size = 200, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.4

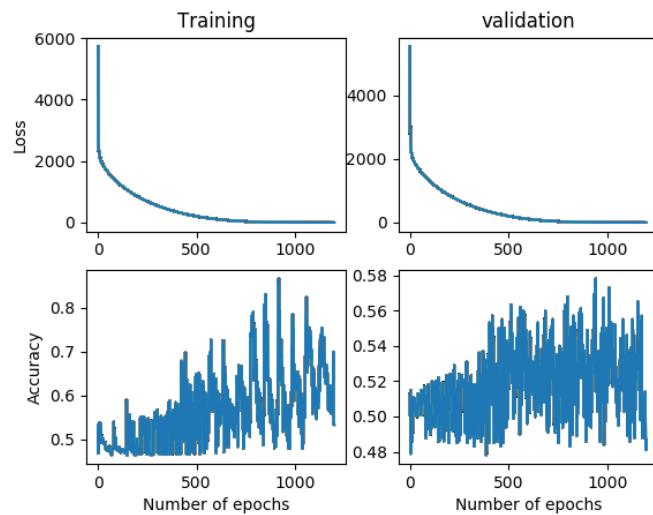


FIGURE 29. Batch Size = 20, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.5

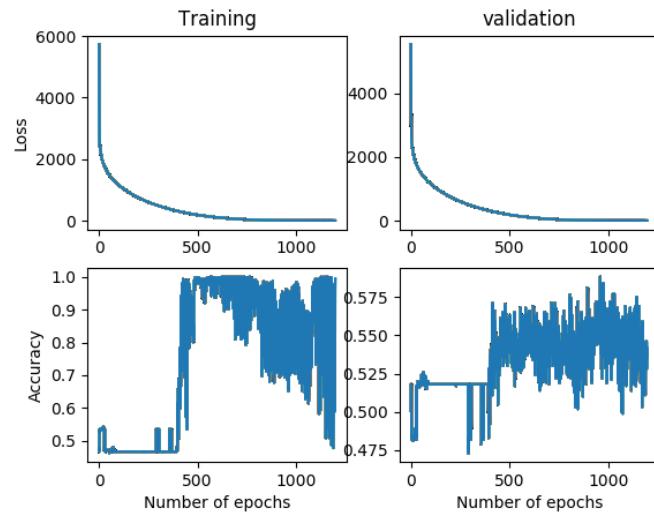


FIGURE 30. Batch Size = 250, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.5

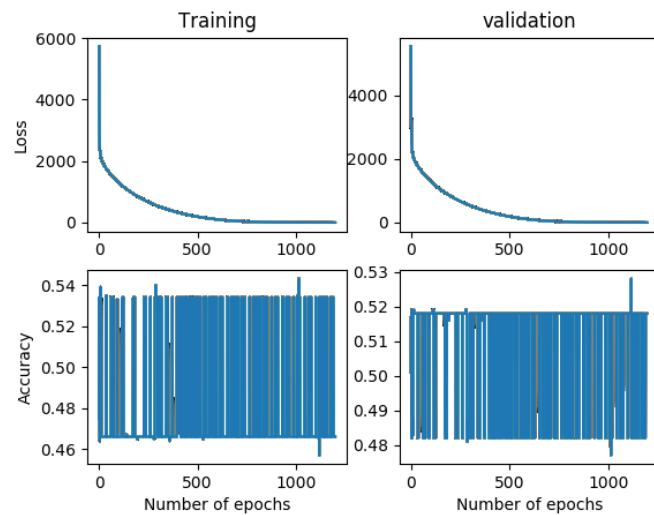


FIGURE 31. Batch Size = 25, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.5

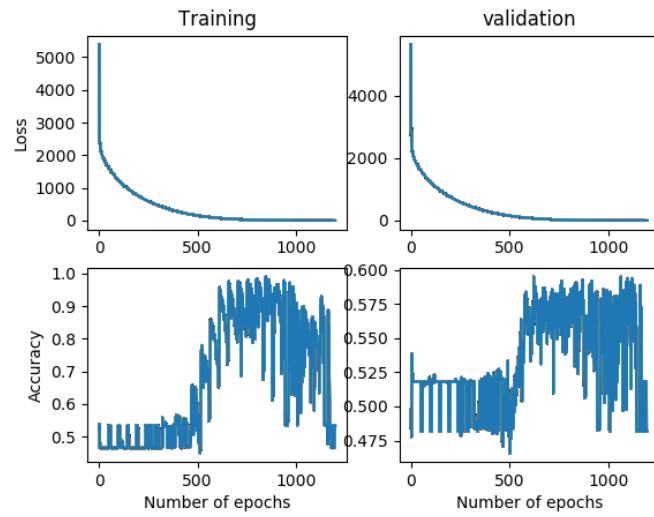


FIGURE 32. Batch Size = 30, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.5

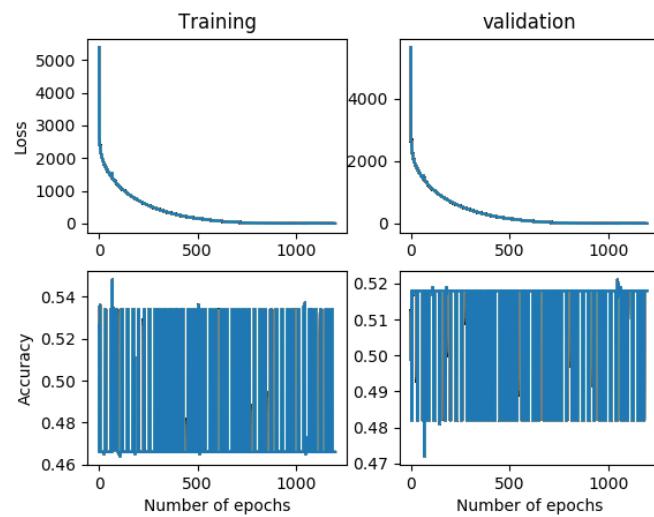


FIGURE 33. Batch Size = 50, Learning Rate = 1e-3, Beta = 0.01, Dropout = 0.5

(10) Visualizations.

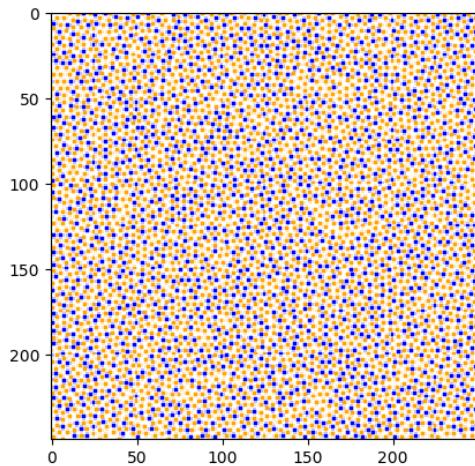


FIGURE 34. Glass original

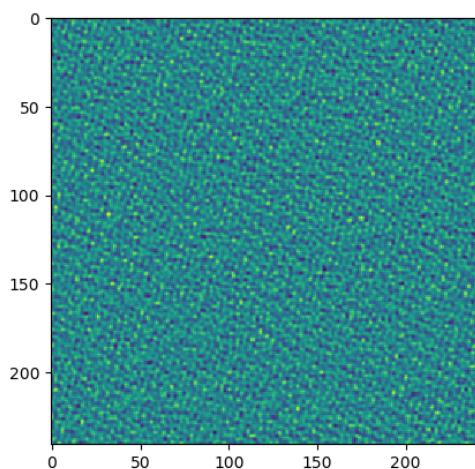


FIGURE 35. Glass channel 1

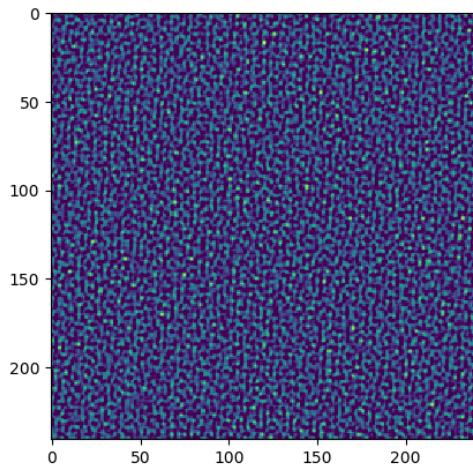


FIGURE 36. Glass channel 2

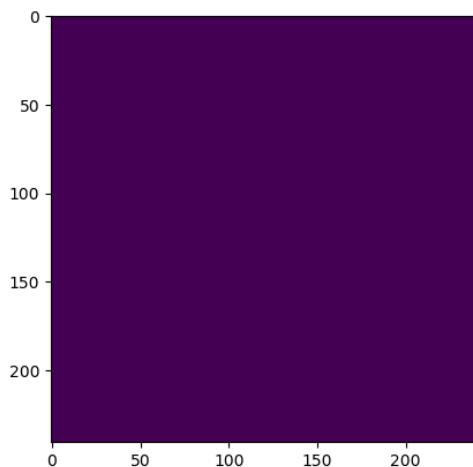


FIGURE 37. Glass channel 3

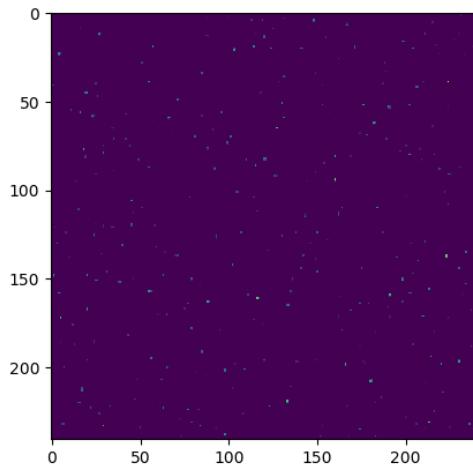


FIGURE 38. Glass channel 4

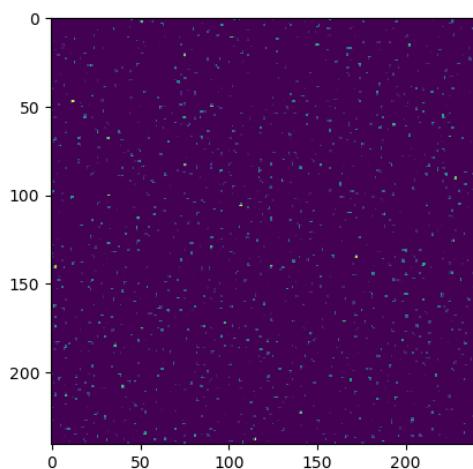


FIGURE 39. Glass channel 5

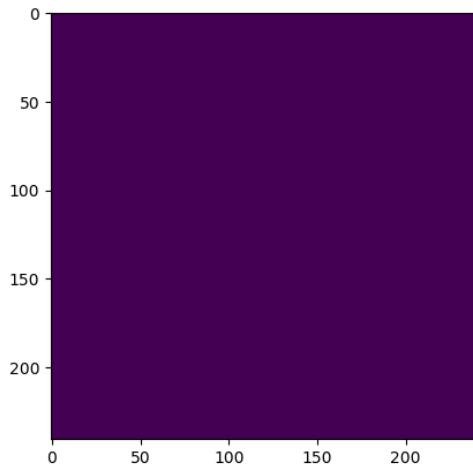


FIGURE 40. Glass channel 6

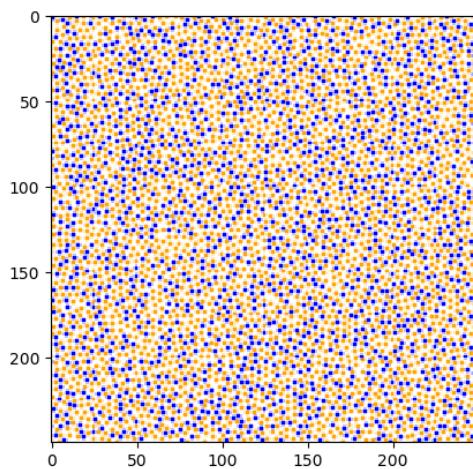


FIGURE 41. Liquid original

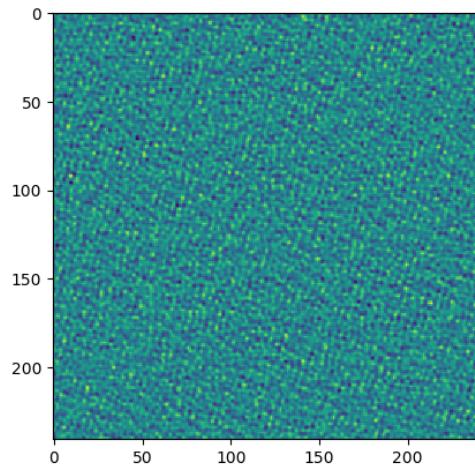


FIGURE 42. Liquid channel 1

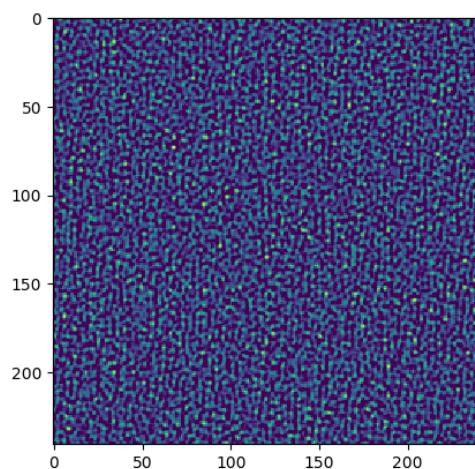


FIGURE 43. Liquid channel 2

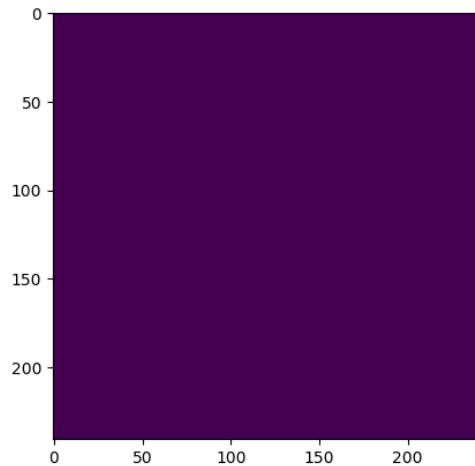


FIGURE 44. Liquid channel 3

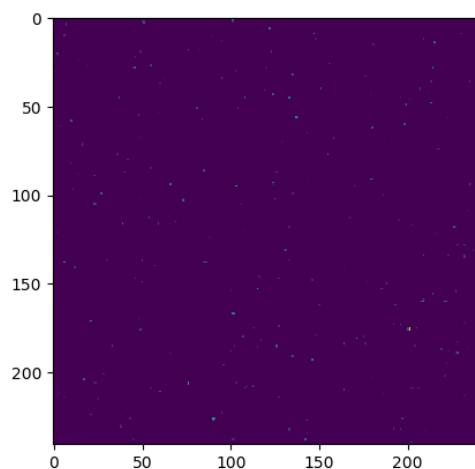


FIGURE 45. Liquid channel 4

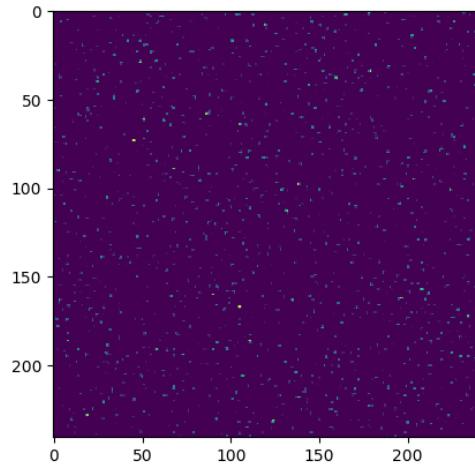


FIGURE 46. Liquid channel 5

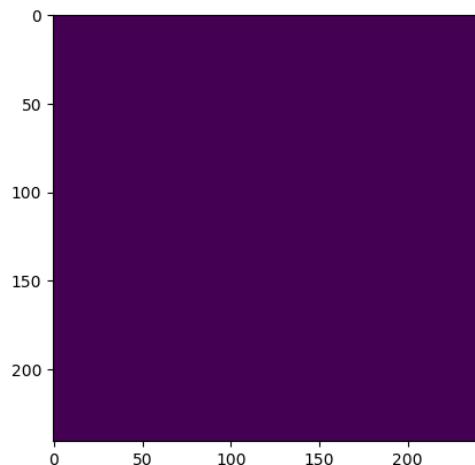


FIGURE 47. Liquid channel 6

2. 12/13/2017

- (1) Updated python scripts and merged with GitHub - the recent additions include streamlining the learning rate schedule and data augmentation options. The following is for a standard run on the original endpoint data:

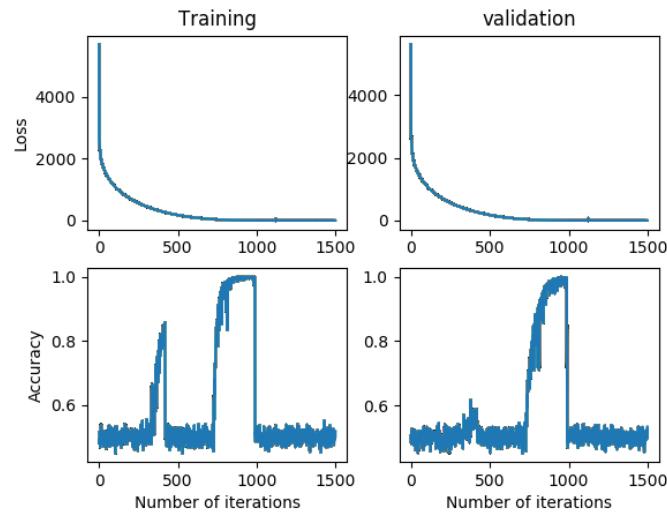


FIGURE 48. Initial Learning Rate = 1e-3, Final Learning Rate = 1e-3, Schedule Length = 550 iterations, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

(2) I then, twice, set the final learning rate to be smaller, at 1e-4:

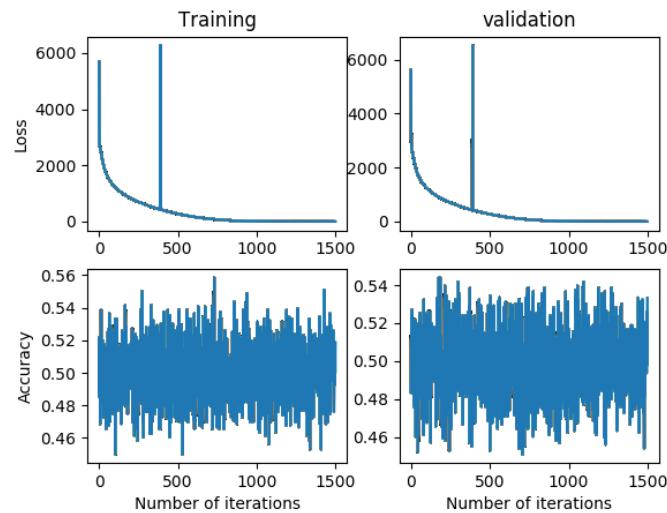


FIGURE 49. Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Schedule Length = 550 iterations, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

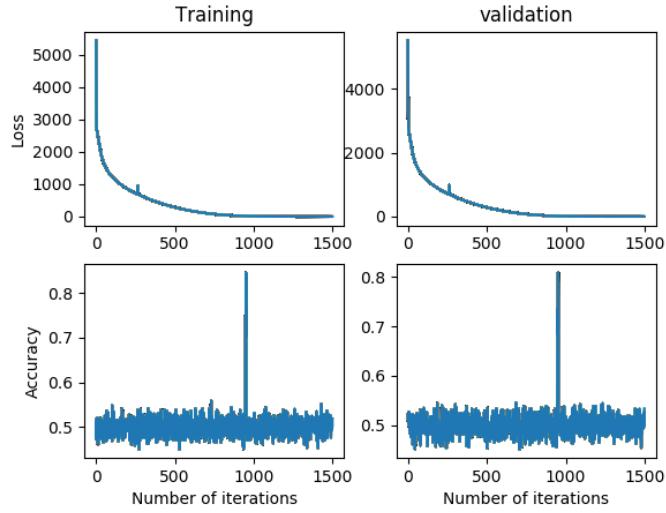


FIGURE 50. Initial Learning Rate =  $1e-3$ , Final Learning Rate =  $1e-3$ , Schedule Length = 550 iterations, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

- (3) What is going on here? It seems as though setting the learning rate a magnitude lower really screws things up...
- (4) Note: one iteration goes through 10 batches, so if the batch size is 80, then one iteration goes through 800 examples. The training set is 70 percent of the total 20,000 images, so the training set is 14,000 images. Thus, one epoch is 17.5 iterations, or approximately 18 iterations. That means that 1500 iterations is about 83 epochs. Question: how many epochs is typically needed for a binary classification problem? And how long does this typically take?
- (5) I also tested the usual hyperparameters on the original endpoint data and turned on the data augmentation, to see if there would be any issues. I got the following training runs:

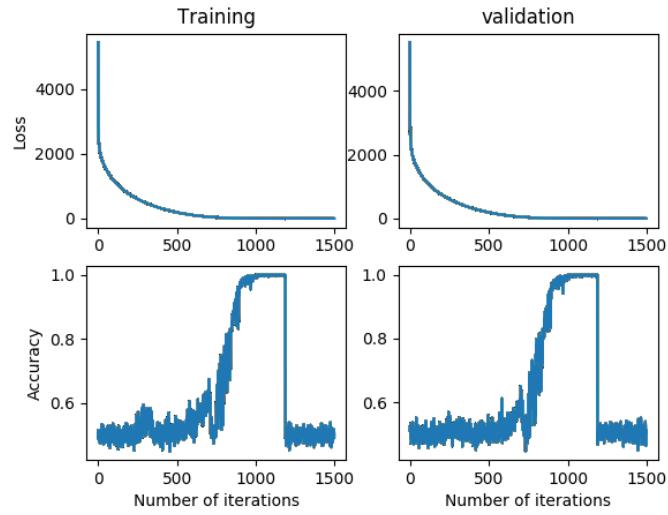


FIGURE 51. Initial Learning Rate = NA, Final Learning Rate = 1e-3, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

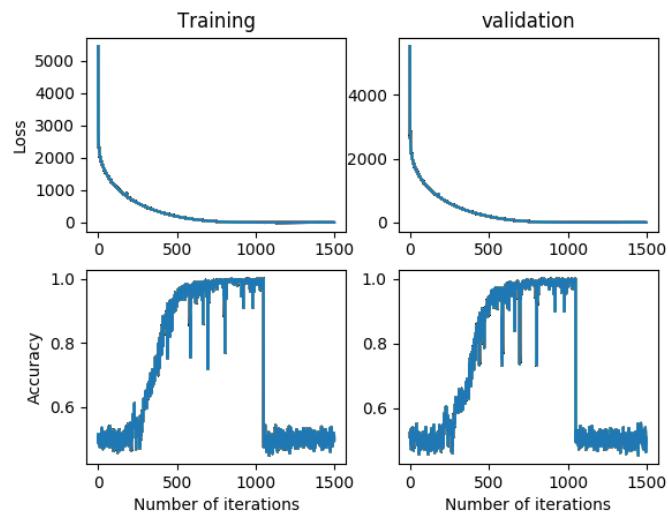


FIGURE 52. Initial Learning Rate = NA, Final Learning Rate = 1e-3, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

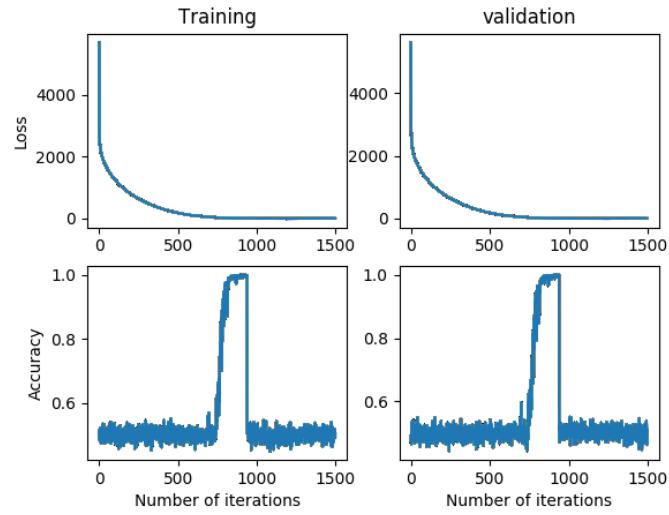


FIGURE 53. Initial Learning Rate = NA, Final Learning Rate = 1e-3, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

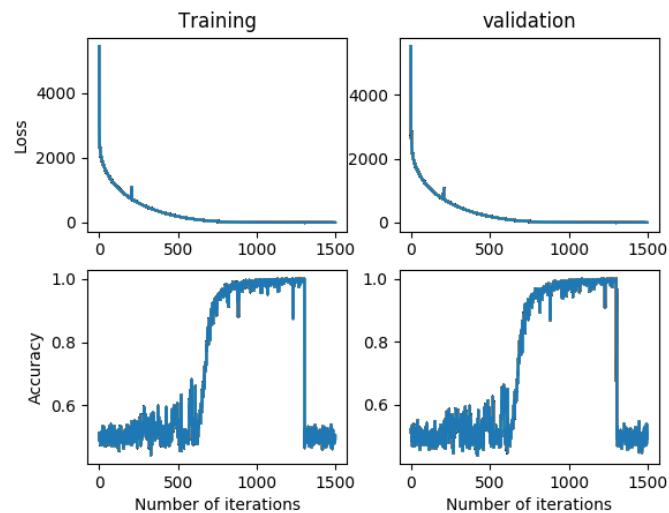


FIGURE 54. Initial Learning Rate = NA, Final Learning Rate = 1e-3, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

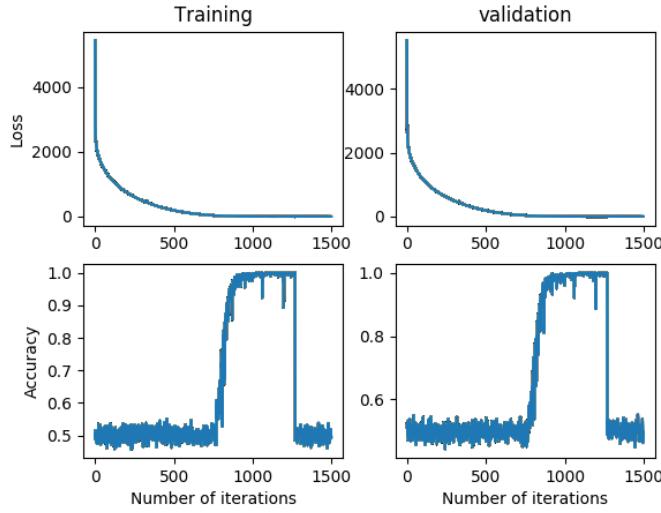


FIGURE 55. Initial Learning Rate = NA, Final Learning Rate = 1e-3, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

(6) I should double check that the dataset augmentation is actually working.

### 3. 12/22/2017

(1) I am now going to run a series of training runs that will explore the relationship between the learning rate schedule and accuracy. First, let's try the threshold idea. Then, we will try the standard exponential decay idea. Let's try the following. When the accuracy gets to 90 percent, let's decrease the learning rate by an order of magnitude. Let's just run that simple test, say, 8 times, and see what happens!

(2) Initial learning rate = , Final learning rate = ,

### 4. 12/24/2017

(1) We ran a series of training runs that used a learning rate threshold. The learning rate starts at the initial learning rate. The learning rate changes to the final learning rate whenever the validation accuracy is above 90 percent. This allows the learning rate to go back to the initial learning rate if the accuracy drops below 90 percent. Here are plots showing six tests. These plots demonstrate that the sharp drops in accuracy were, indeed, the result of a learning rate that was too large towards the end of training, since there are no sharp drops in accuracy here:

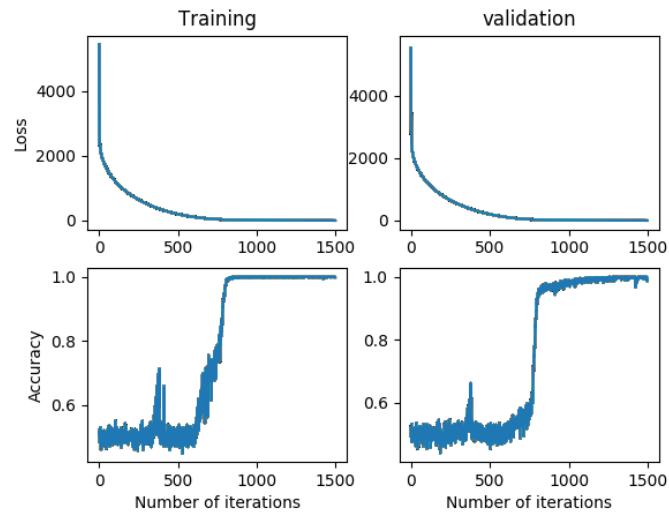


FIGURE 56. Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

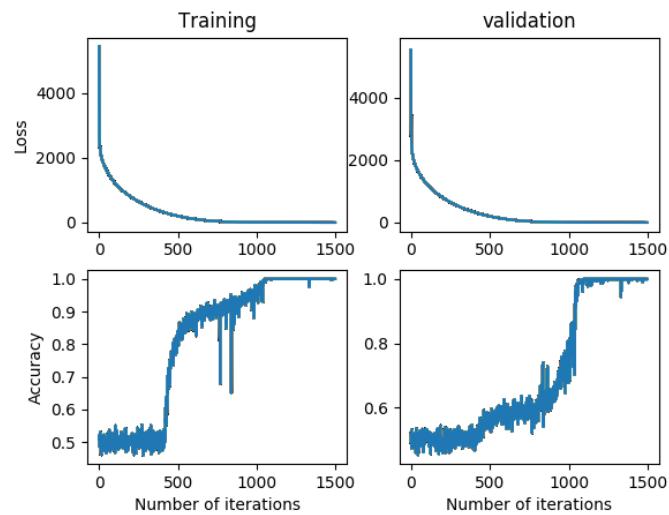


FIGURE 57. Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

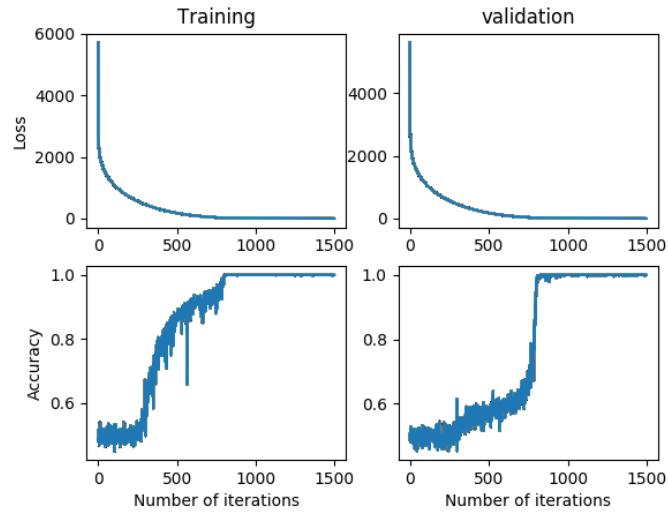


FIGURE 58. Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

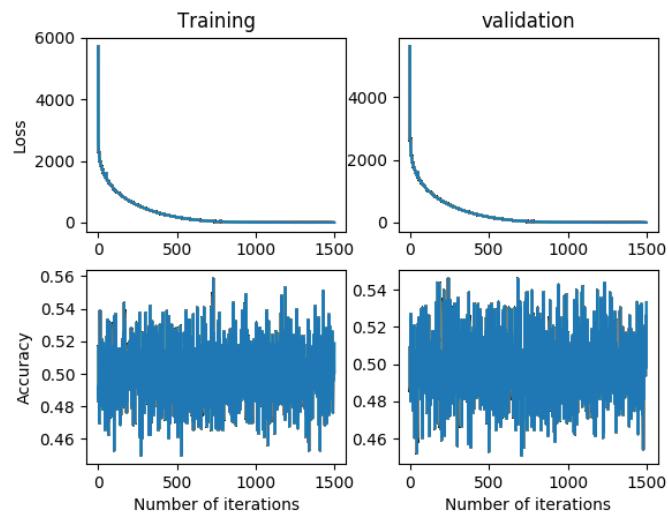


FIGURE 59. Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

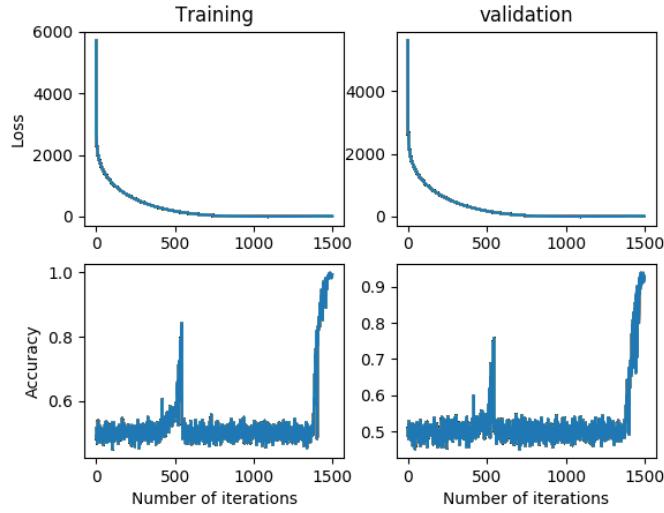


FIGURE 60. Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

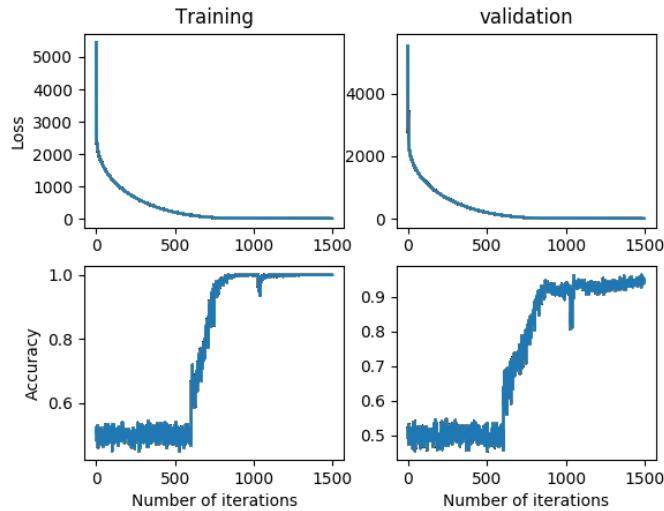


FIGURE 61. Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

#### 4.1. 12/29/2017.

- (1) Ran another series of runs that used a similar learning rate threshold just to check that my GitHub had been updated appropriately. This time used only 1e-3 to 1e-4. In one example, the accuracy crashed. Perhaps evidences that smaller learning rates are needed closer to the minimum. For the sake of completeness, these extra runs are reported below:

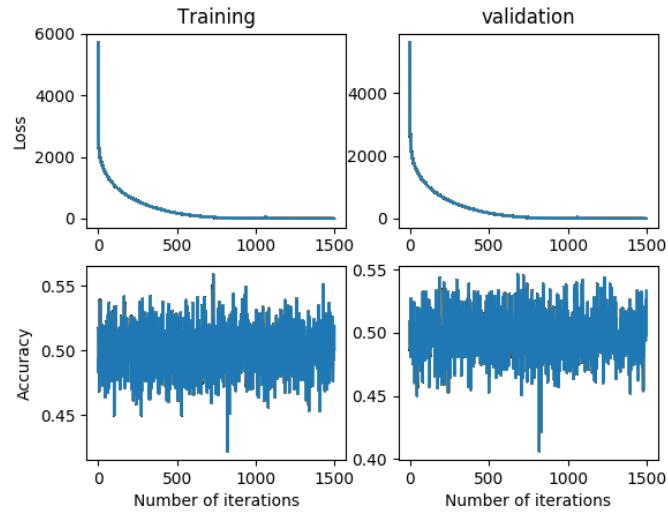


FIGURE 62. Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

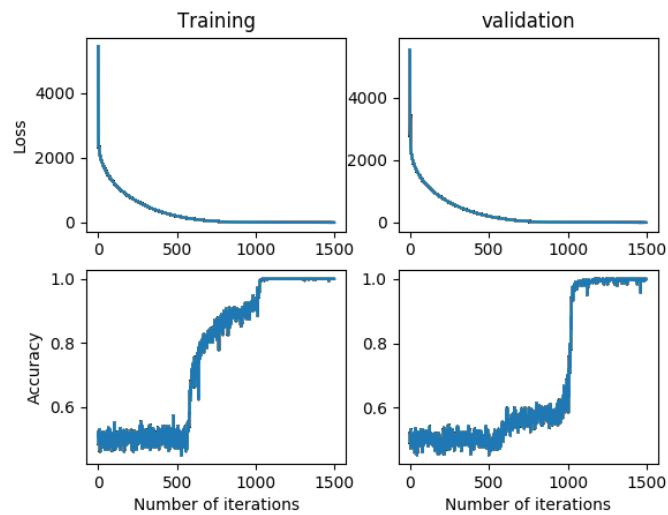


FIGURE 63. Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

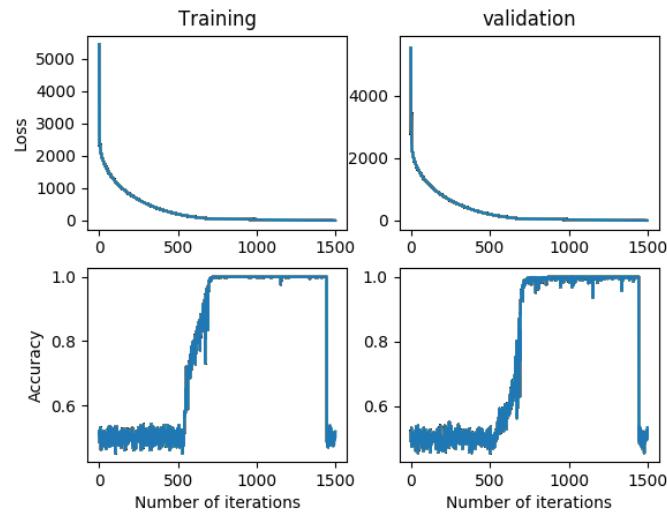


FIGURE 64. Initial Learning Rate =  $1e-3$ , Final Learning Rate =  $1e-4$ , Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

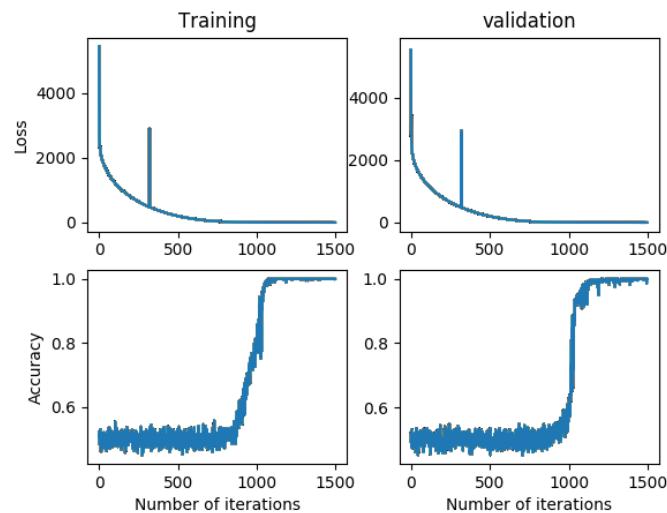


FIGURE 65. Initial Learning Rate =  $1e-3$ , Final Learning Rate =  $1e-4$ , Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

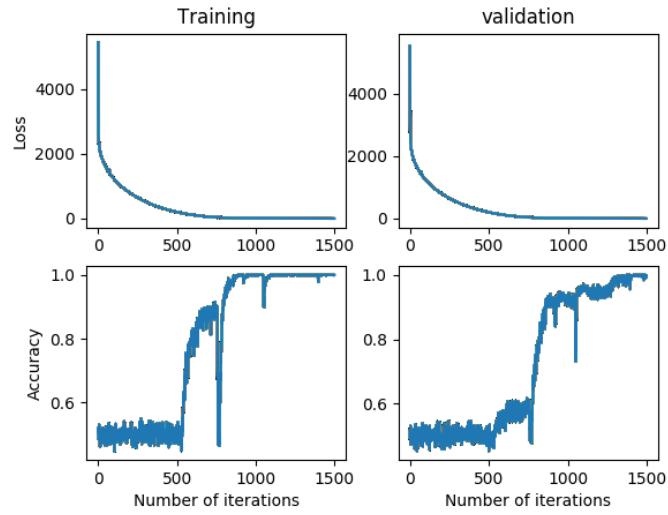


FIGURE 66. Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

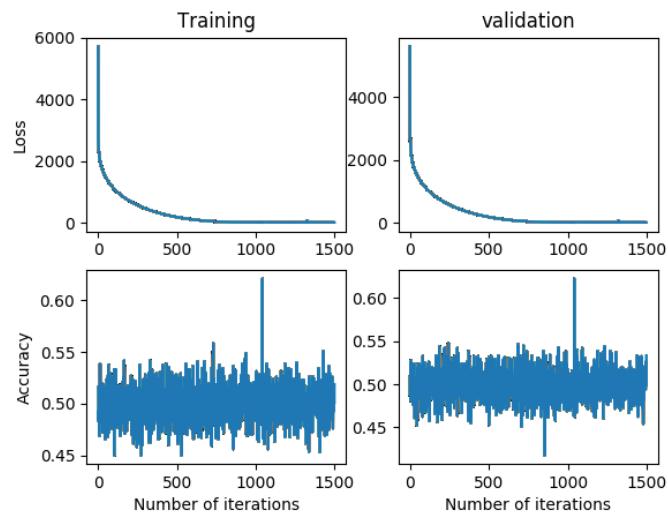


FIGURE 67. Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

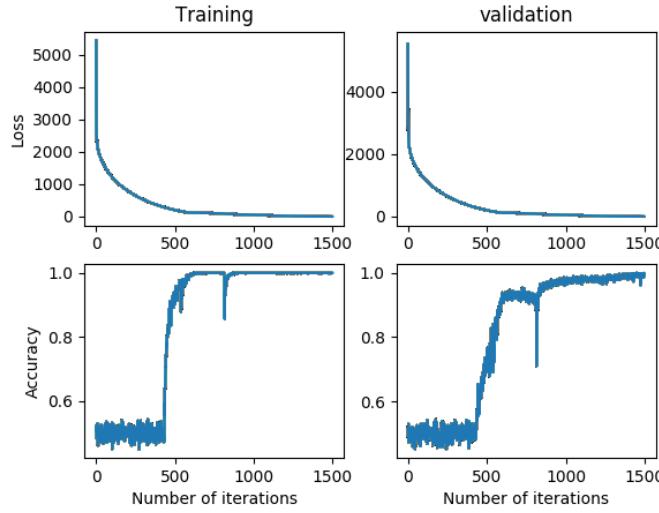


FIGURE 68. Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5

#### 4.2. 12/30/2017.

- (1) Tested the data augmentation scheme manually. Printed out a series of images and their respective augmented versions that are given when generator.next is invoked, respectively with the data aug option set to false and with the data aug option set to true.

#### 4.3. 1/1/2018.

- (1) Added main\_glassliquid\_restart.py, a file that can run training starting from a previous saved model. Tested it. Now, let's try to train again on the dataset that consists of glasses and liquids that are each 12 time steps away from the assumed glass transition temperature of 0.21. Previously, the best parameters were batch size 200, learning rate 1e-3, beta 0.01, dropout 0.4. Now, we are going to do a series of simulations using the same parameters, except, first, we are going to include dataset augmentation, and second, we are going to run the simulation at least twice as long. Then, once we get a sense for how these train, we will add in a learning rate threshold step, or even multiple threshold steps if necessary!

#### 4.4. 1/4/2018.

- (1) Realized that I had run on the original endpoint data. So, copied over the folder called "data\_12steps" into the current machine-learning-glasses directory, removed the current metadata folder, and created a new metadata folder with this data. Then, set up new training runs again.

### 5. 1/9/2018

- (1) Collected data from the above-mentioned simulations. The training curves are as follows:

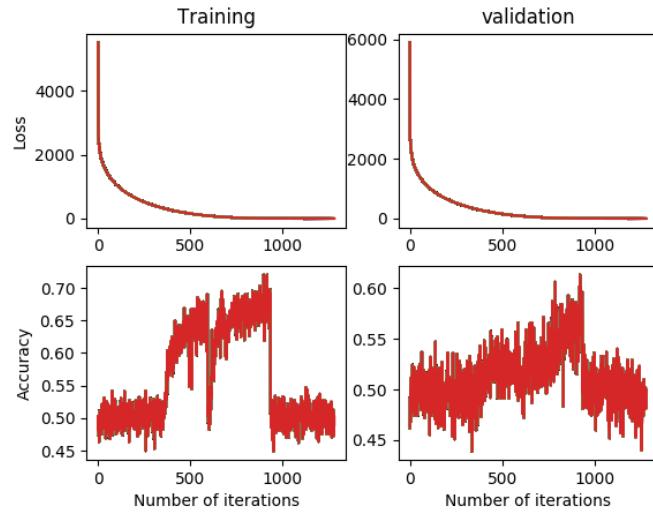


FIGURE 69. Version 1: Initial Learning Rate =  $1e-3$ , Final Learning Rate =  $1e-4$ , Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.4, data aug = true

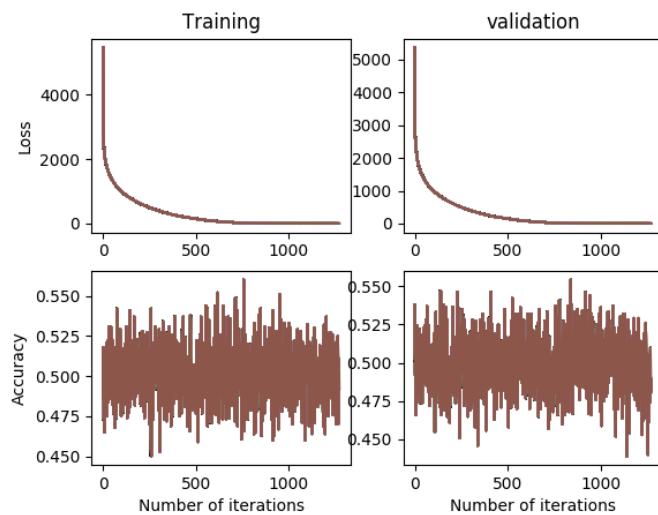


FIGURE 70. Version 2: Initial Learning Rate =  $1e-3$ , Final Learning Rate =  $1e-4$ , Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.4, data aug = true

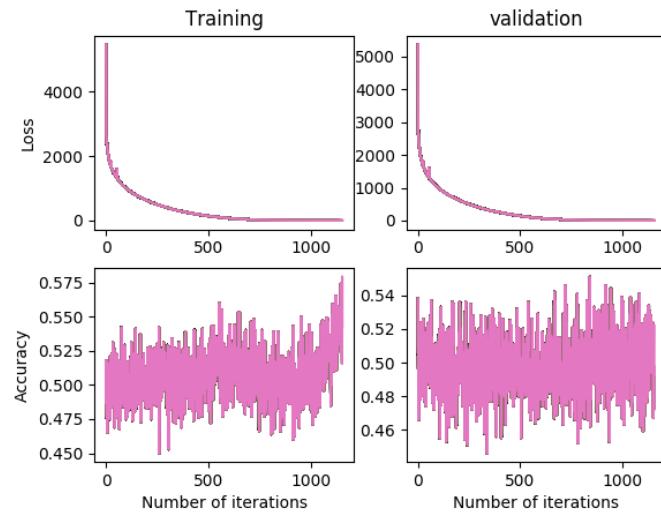


FIGURE 71. Version 3: Initial Learning Rate =  $1e-3$ , Final Learning Rate =  $1e-4$ , Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.4, data aug = true

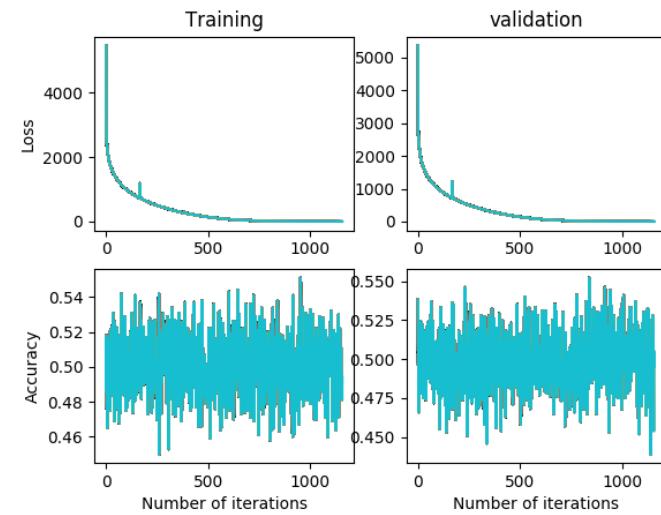


FIGURE 72. Version 4: Initial Learning Rate =  $1e-3$ , Final Learning Rate =  $1e-4$ , Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.4, data aug = true

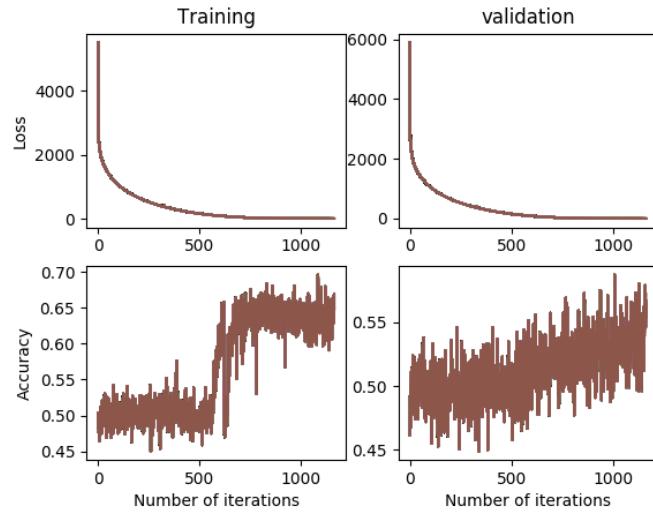


FIGURE 73. Version 5: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.4, data aug = true

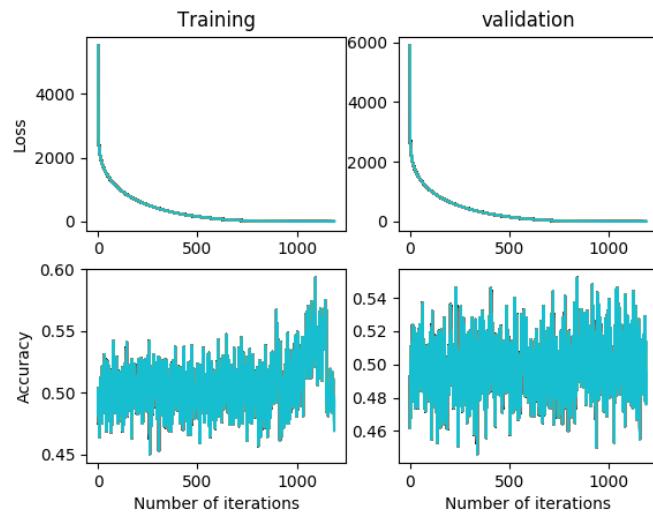


FIGURE 74. Version 6: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.4, data aug = true

- (2) The problem is that my allotted time (36 hours) was not enough, so the simulations terminated early. My program, however, was only set to save the current model at the end of the iterations. So, the current model was not saved, only the best model was saved. So, I went back into the main code and changed it so that the current model is updated and saved at every iteration, in addition to the best

model being updated when necessary. I also included restart capability into the main code (using parsing options), so that there is not a separate restart file.

- (3) I am now running 7 new simulations. The first is just a re-do of version1 step 0. The others take version5 step 0 and start from the saved best model as a starting point, using the new restart capabilities of the main code. The first two simulations from this starting point (version5 step2 and version5 step2 trial2) use the original parameters. The next two (version5 step2 trial3 and version5 step2 trial4) use  $\eta_{\text{initial}}=1e-4$  and  $\eta_{\text{final}}=1e-5$ , and the final two (version5 step2 trial5 and version5 step2 trial6) use  $\eta_{\text{initial}}=1e-5$  and  $\eta_{\text{final}}=1e-6$ . We will see how these 7 new runs turn out.

### 5.1. 1/14/2018.

- (1) Here is the result for the continuations from Version 5 step 0, i.e. Version 5 step 1 trials 1 through 6.

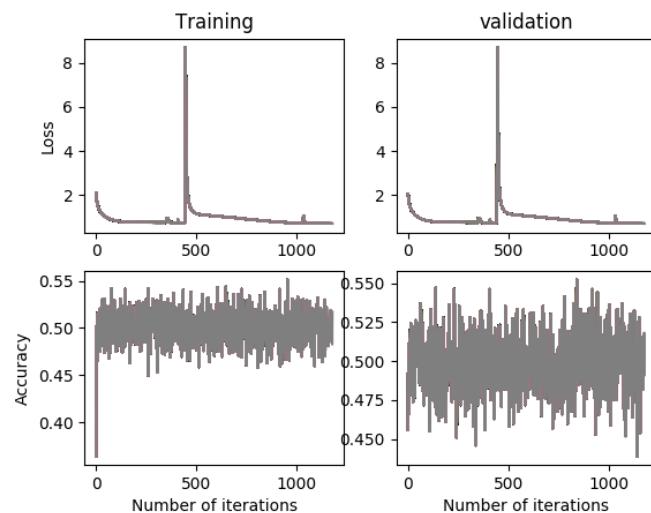


FIGURE 75. Version 5 step 1 trial 1: Initial Learning Rate =  $1e-3$ , Final Learning Rate =  $1e-4$ , Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.4, data aug = true

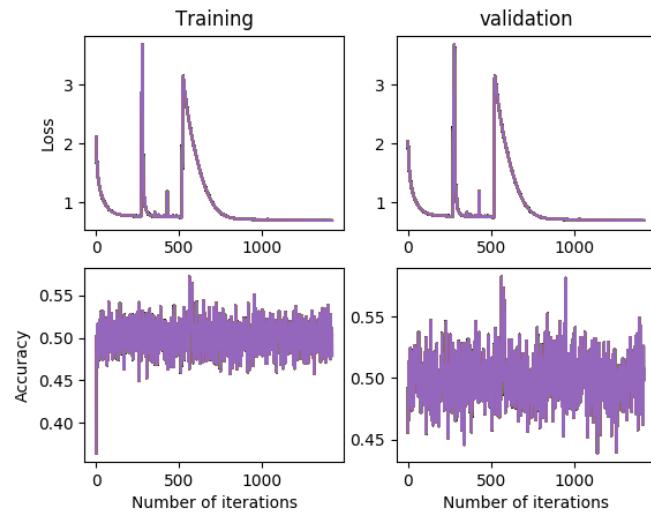


FIGURE 76. Version 5 step 1 trial 2: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.4, data aug = true

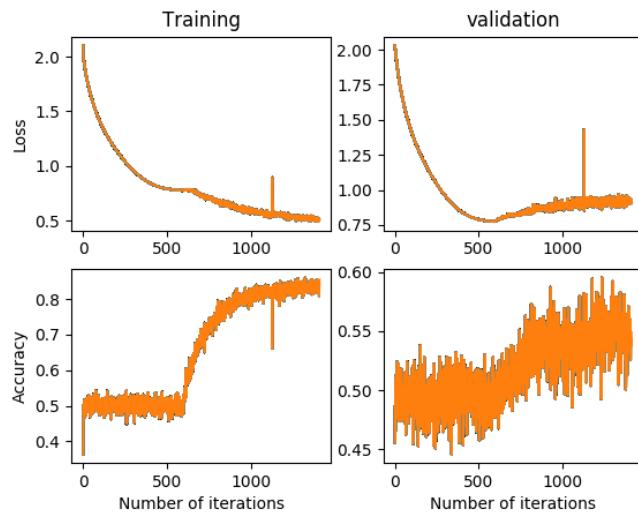


FIGURE 77. Version 5 step 1 trial 3: Initial Learning Rate = 1e-4, Final Learning Rate = 1e-5, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.4, data aug = true

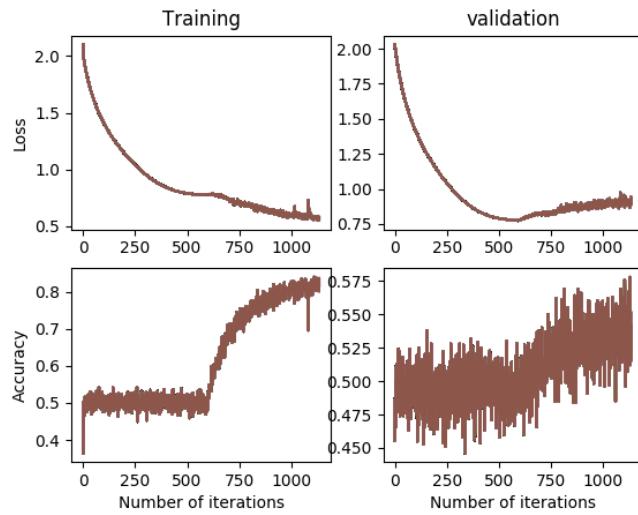


FIGURE 78. Version 5 step 1 trial 4: Initial Learning Rate = 1e-4, Final Learning Rate = 1e-5, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.4, data aug = true

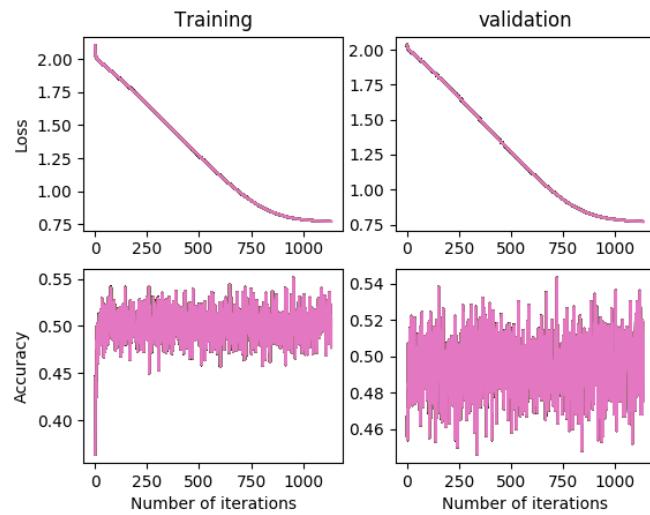


FIGURE 79. Version 5 step 1 trial 5: Initial Learning Rate = 1e-5, Final Learning Rate = 1e-6, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.4, data aug = true

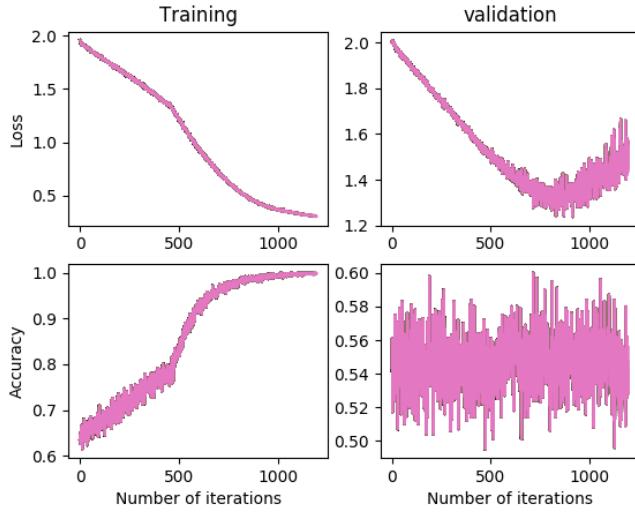


FIGURE 80. Version 5 step 1 trial 6: Initial Learning Rate = 1e-5, Final Learning Rate = 1e-6, Schedule Length = NA, Batch Size = 80, L2 beta = 0.01, Dropout = 0.4, data aug = true

(2) I am now going to generate new data that is farther away from the assumed glass transition temperature of  $T_g = 0.21$ . Based on cooling.dat file in the folder N2e7/0, which is the first of 10,000 simulations, this temperature is at the 184th line of the file (0.21575), which is the 183rd time listed. There are 200 times listed in cooling.dat, each one separated by a time length of 100,000 time steps. This is because the total simulation is  $2e7$ , or 20 million. In the folder /project/depablo/kswanson/2017, I open the file make\_images\_kirk.sh and change the liquid generation number to 500000 and keep the glass generation number at 20000000.

## 5.2. 1/15/2018.

- (1) After generating the new datapoint, I create a folder called data\_liquid5 and transfer it to the git repo. Now, I delete the current metadata folder and create a new metadata folder that contains the data\_liquid5 in it.
- (2) V1-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 10, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (3) V2-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 10, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (4) V3-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 20, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (5) V4-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 20, L2 beta = 0.01, Dropout = 0.5, data aug = true.

- (6) V5-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 30, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (7) V6-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 30, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (8) V7-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 50, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (9) V8-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 50, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (10) V9-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (11) V10-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5, data aug = true.

### 5.3. 1/19/2018.

- (1) Using timestep 5 (500000) for the liquid and timestep 200 (20000000) for the glass, here are training results:

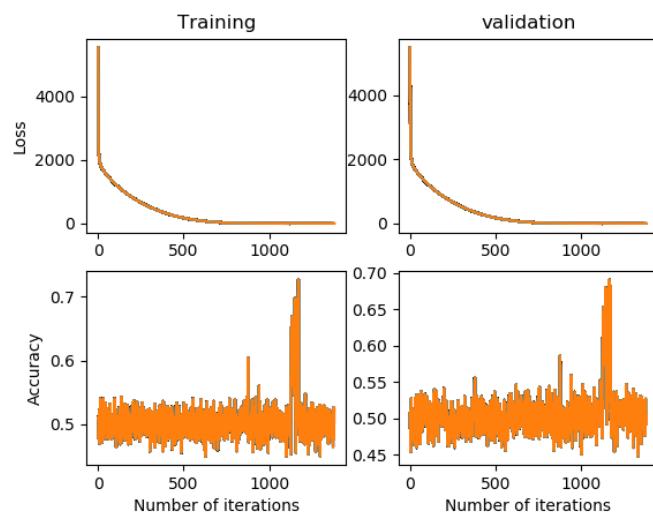


FIGURE 81. Version 1 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 10, L2 beta = 0.01, Dropout = 0.5, data aug = true

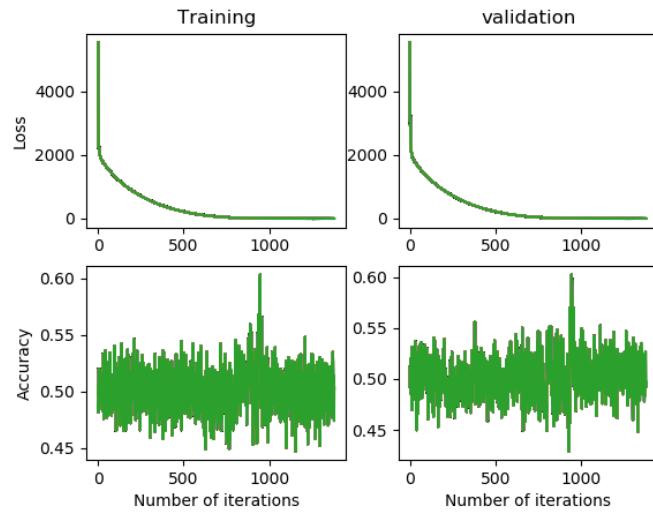


FIGURE 82. Version 2 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 10, L2 beta = 0.01, Dropout = 0.5, data aug = true

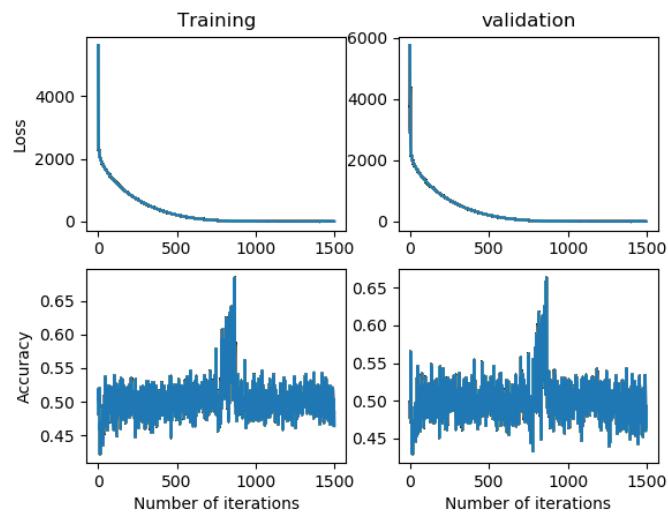


FIGURE 83. Version 3 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 20, L2 beta = 0.01, Dropout = 0.5, data aug = true

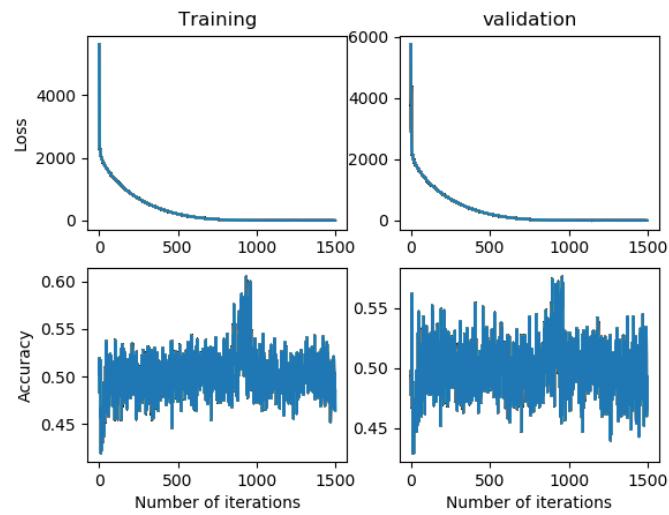


FIGURE 84. Version 4 step 1: Initial Learning Rate =  $1e-3$ , Final Learning Rate =  $1e-4$ , Eta threshold = 0.90, Batch Size = 20, L2 beta = 0.01, Dropout = 0.5, data aug = true

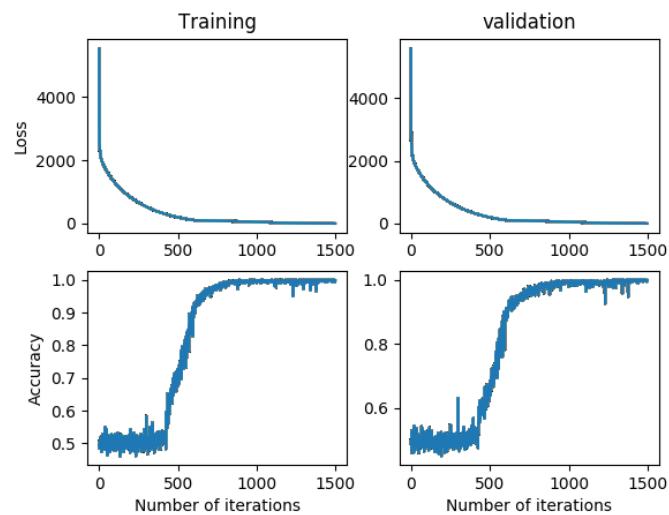


FIGURE 85. Version 5 step 1: Initial Learning Rate =  $1e-3$ , Final Learning Rate =  $1e-4$ , Eta threshold = 0.90, Batch Size = 30, L2 beta = 0.01, Dropout = 0.5, data aug = true

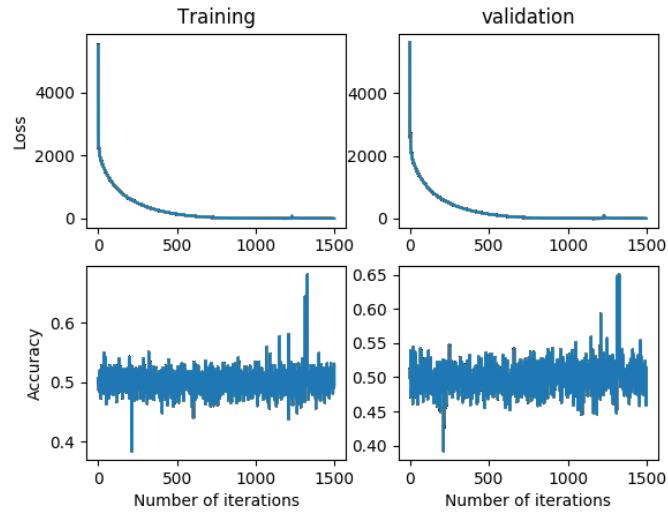


FIGURE 86. Version 6 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 30, L2 beta = 0.01, Dropout = 0.5, data aug = true

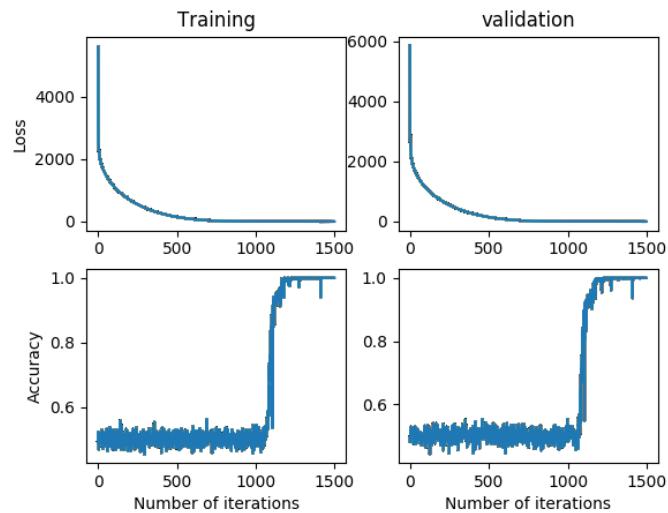


FIGURE 87. Version 7 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 50, L2 beta = 0.01, Dropout = 0.5, data aug = true

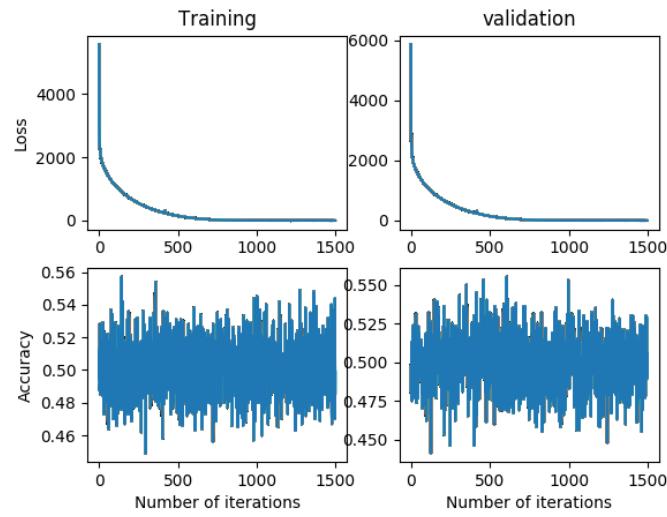


FIGURE 88. Version 8 step 1: Initial Learning Rate =  $1e-3$ , Final Learning Rate =  $1e-4$ , Eta threshold = 0.90, Batch Size = 50, L2 beta = 0.01, Dropout = 0.5, data aug = true

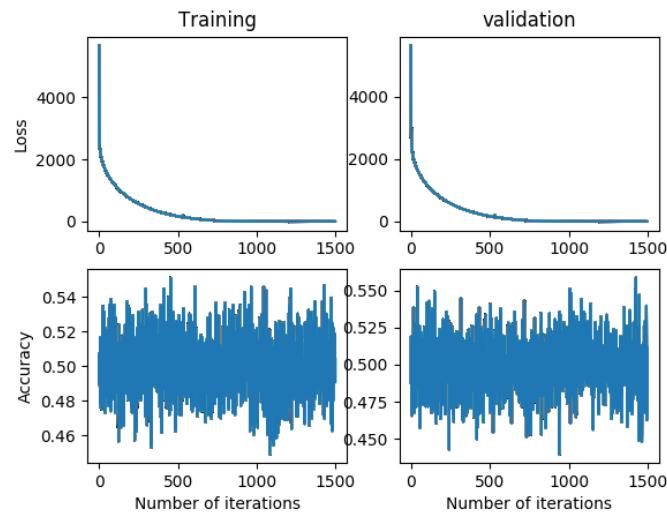


FIGURE 89. Version 9 step 1: Initial Learning Rate =  $1e-3$ , Final Learning Rate =  $1e-4$ , Eta threshold = 0.90, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5, data aug = true

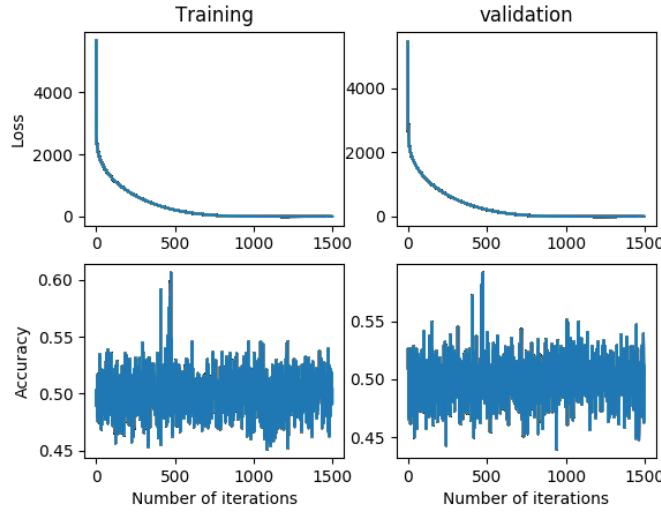


FIGURE 90. Version 10 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5, data aug = true

- (2) Version 5 Step 1 had a final validation accuracy of 99.7% and a final test accuracy of 99.7333%. Added changes to main\_glassliquid.py so that by setting -restart=True and -evaluation=True in the options, and setting -restart\_file to the desired filename, one can check the final validation and test accuracy of a given model.
- (3) Version 7 Step 1 had a final validation accuracy of 99.8% and a final test accuracy of 99.8333%.
- (4) Now I am generating images for liquid timestep 10 (1000000).

#### 5.4. 1/20/2018.

- (1) I am now switching metadata to be from data\_liquid10, our new current dataset.
- (2) V1-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 10, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (3) V2-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 10, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (4) V3-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 20, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (5) V4-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 20, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (6) V5-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 30, L2 beta = 0.01, Dropout = 0.5, data aug = true.

- (7) V6-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 30, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (8) V7-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 50, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (9) V8-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 50, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (10) V9-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (11) V10-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (12) V11-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 100, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (13) V12-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 100, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (14) V13-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 130, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (15) V14-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 130, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (16) V15-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 150, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (17) V16-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 150, L2 beta = 0.01, Dropout = 0.5, data aug = true.

## 5.5. 1/23/2018.

- (1) Collected results from training runs on the data\_liquid10 data. Here they are:

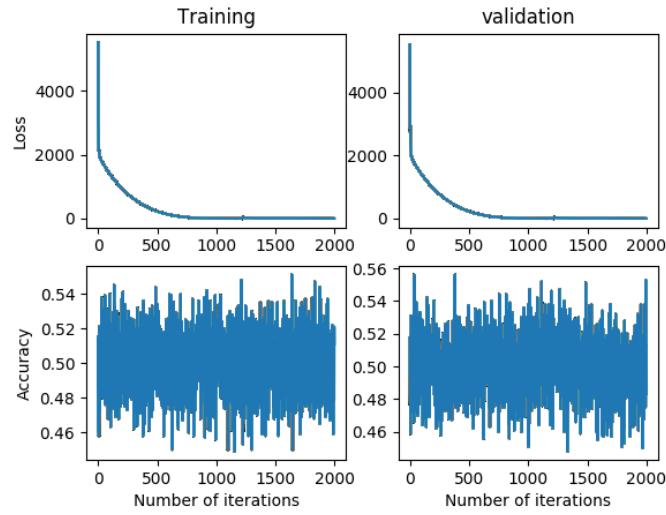


FIGURE 91. Version 1 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 10, L2 beta = 0.01, Dropout = 0.5, data aug = true

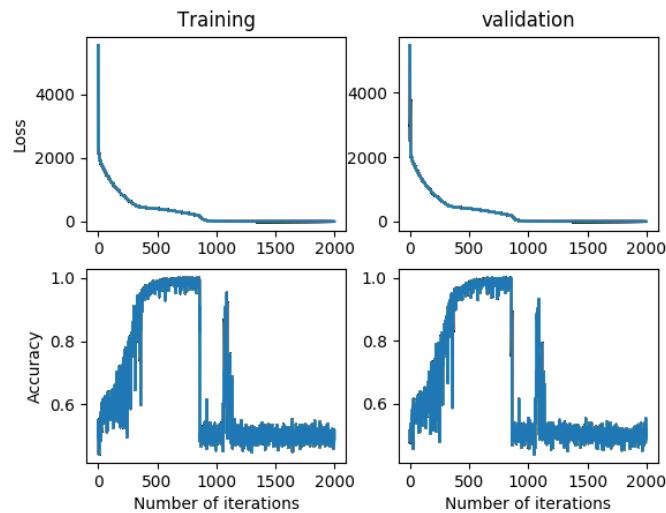


FIGURE 92. Version 2 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 10, L2 beta = 0.01, Dropout = 0.5, data aug = true

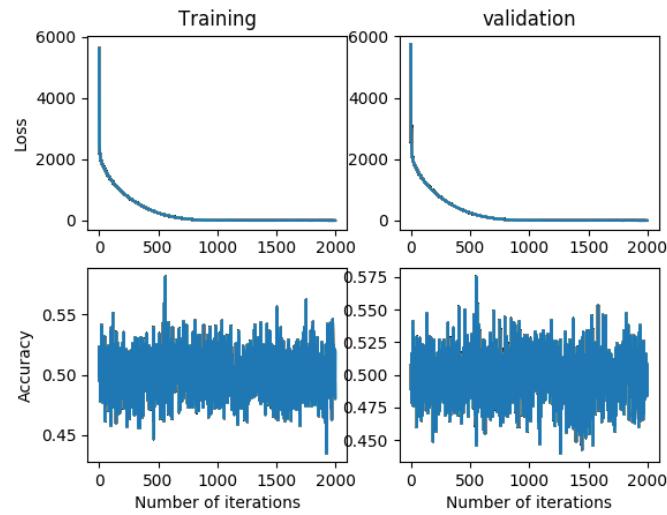


FIGURE 93. Version 3 step 1: Initial Learning Rate =  $1e-3$ , Final Learning Rate =  $1e-4$ , Eta threshold = 0.90, Batch Size = 20, L2 beta = 0.01, Dropout = 0.5, data aug = true

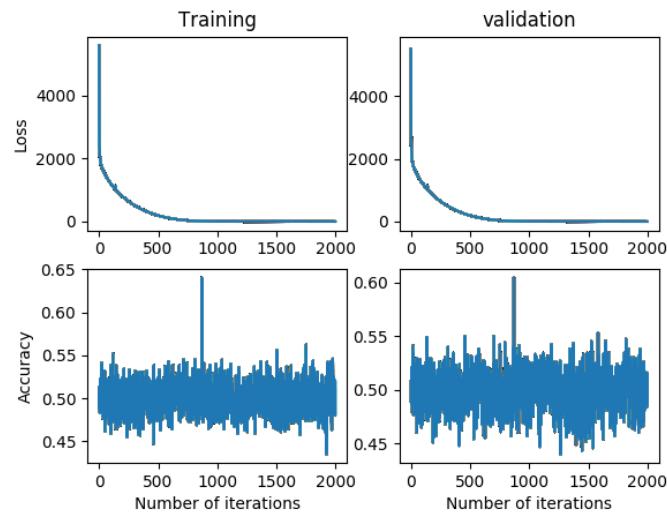


FIGURE 94. Version 4 step 1: Initial Learning Rate =  $1e-3$ , Final Learning Rate =  $1e-4$ , Eta threshold = 0.90, Batch Size = 20, L2 beta = 0.01, Dropout = 0.5, data aug = true

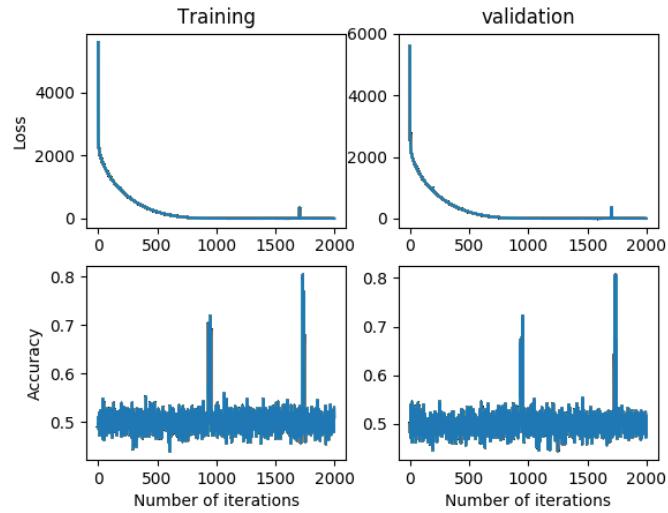


FIGURE 95. Version 5 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 30, L2 beta = 0.01, Dropout = 0.5, data aug = true

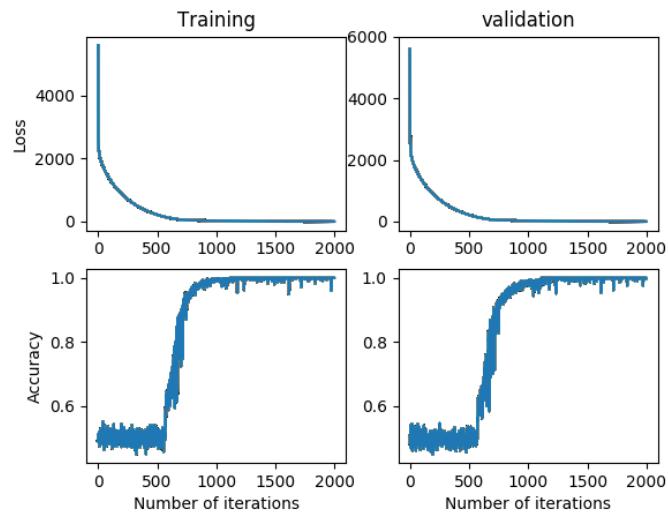


FIGURE 96. Version 6 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 30, L2 beta = 0.01, Dropout = 0.5, data aug = true

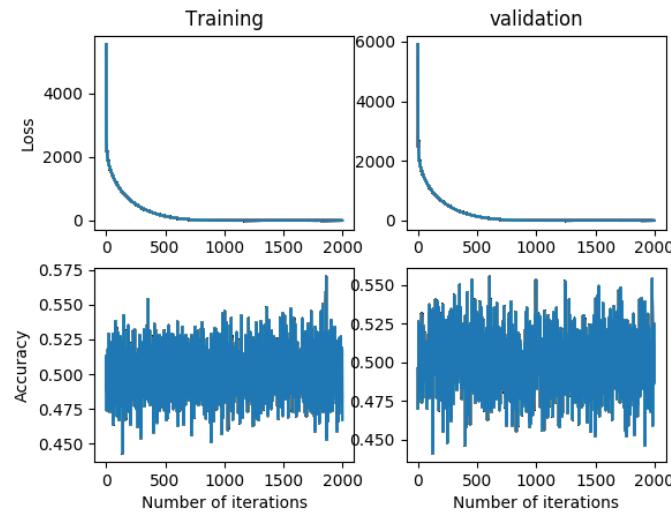


FIGURE 97. Version 7 step 1: Initial Learning Rate =  $1e-3$ , Final Learning Rate =  $1e-4$ , Eta threshold = 0.90, Batch Size = 50, L2 beta = 0.01, Dropout = 0.5, data aug = true

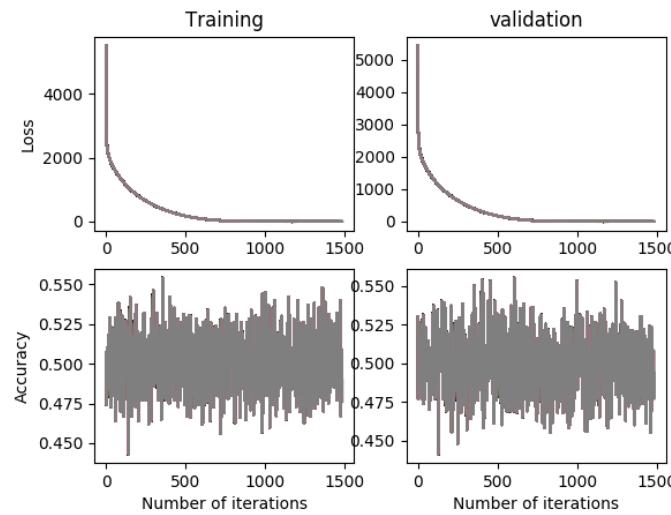


FIGURE 98. Version 8 step 1: Initial Learning Rate =  $1e-3$ , Final Learning Rate =  $1e-4$ , Eta threshold = 0.90, Batch Size = 50, L2 beta = 0.01, Dropout = 0.5, data aug = true

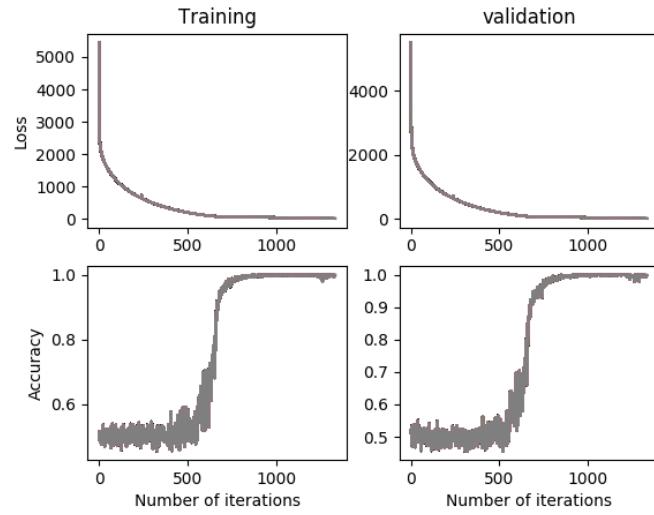


FIGURE 99. Version 9 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5, data aug = true

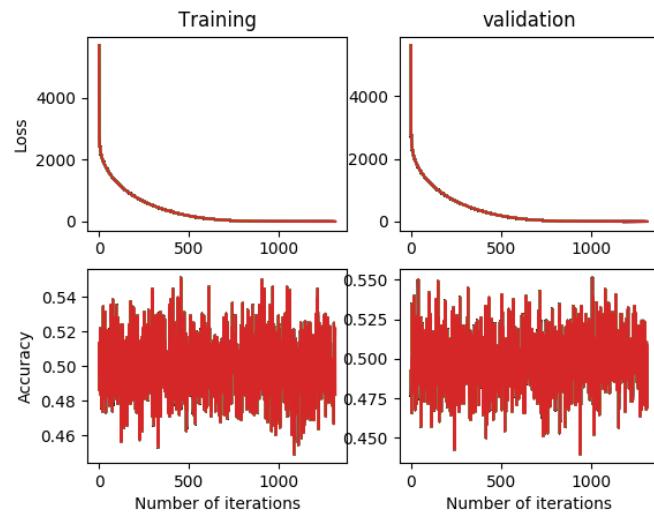


FIGURE 100. Version 10 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5, data aug = true

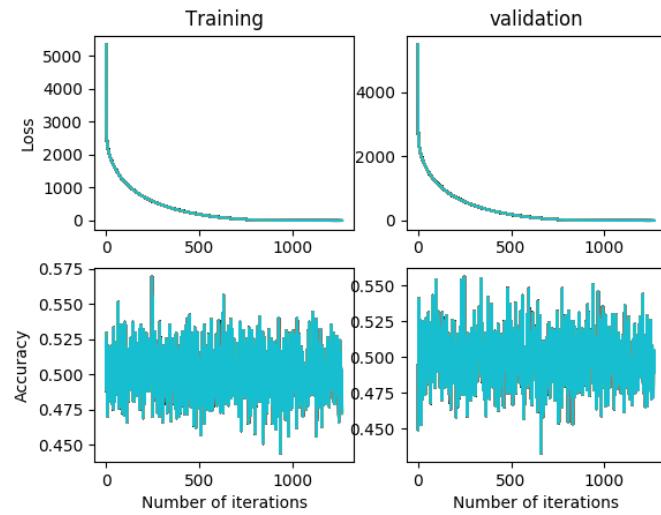


FIGURE 101. Version 11 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 100, L2 beta = 0.01, Dropout = 0.5, data aug = true

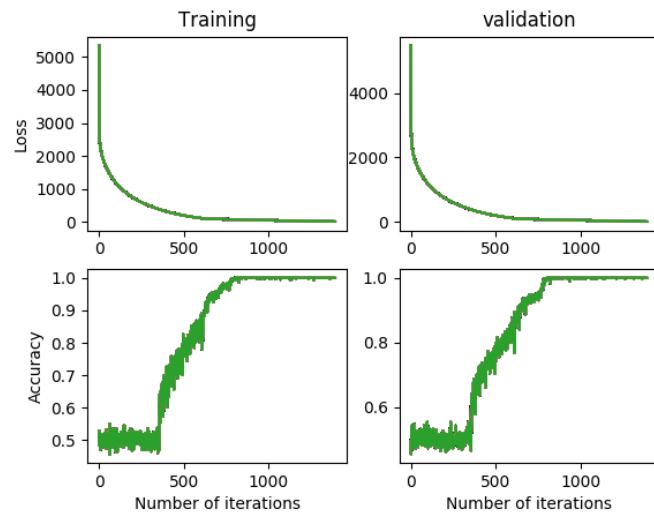


FIGURE 102. Version 12 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 100, L2 beta = 0.01, Dropout = 0.5, data aug = true

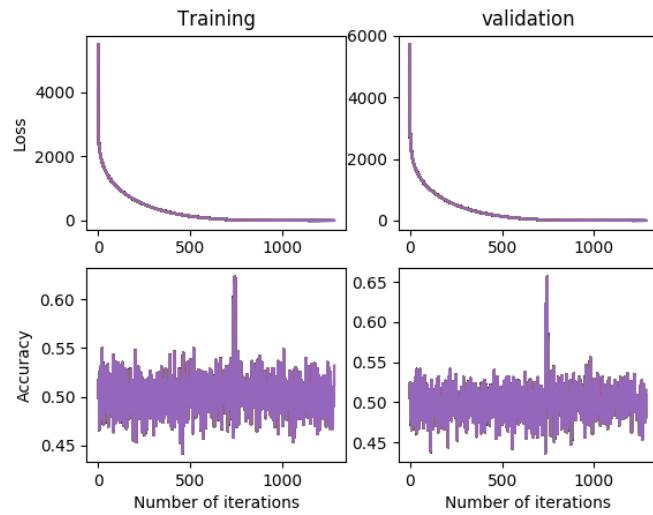


FIGURE 103. Version 13 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 130, L2 beta = 0.01, Dropout = 0.5, data aug = true

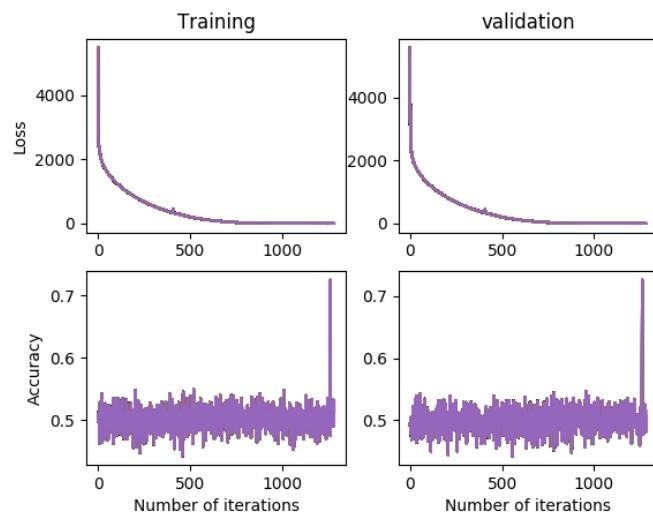


FIGURE 104. Version 14 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 130, L2 beta = 0.01, Dropout = 0.5, data aug = true

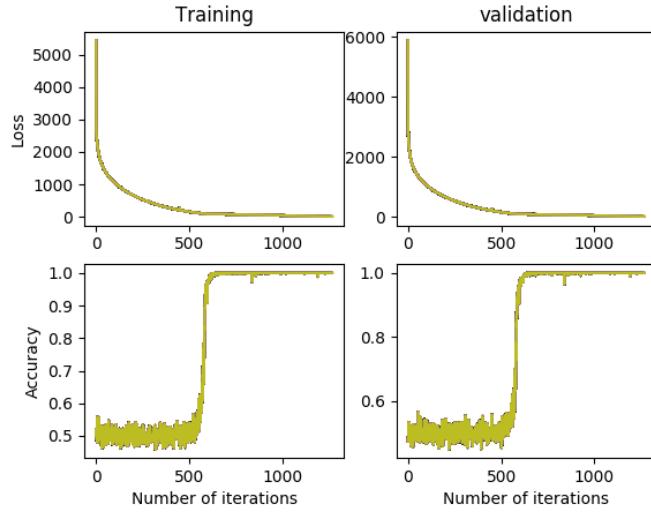


FIGURE 105. Version 15 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 150, L2 beta = 0.01, Dropout = 0.5, data aug = true

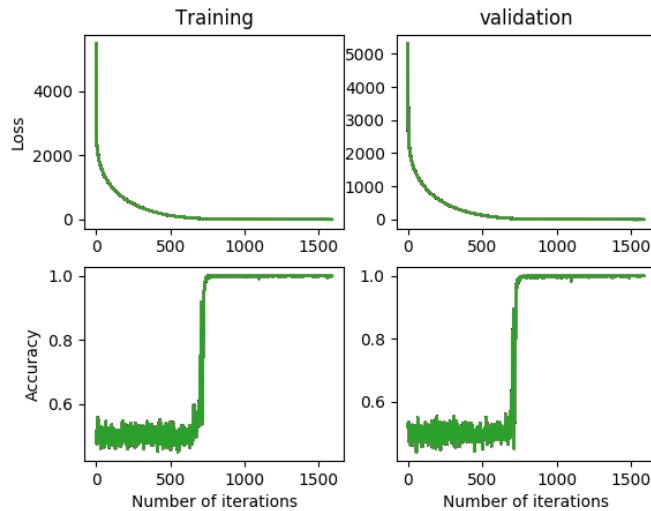


FIGURE 106. Version 16 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 150, L2 beta = 0.01, Dropout = 0.5, data aug = true

- (2) Version 16 Step 1 best model had a final validation accuracy of 99.9% and a final test accuracy of 99.9333%.
- (3) Now I am going to save this model, and delete the others.
- (4) Moved the data\_liquid15 file to the working directory. Deleting metadata and creating a new metadata for this dataset.

- (5) V1-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 10, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (6) V2-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 10, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (7) V3-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 20, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (8) V4-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 20, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (9) V5-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 30, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (10) V6-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 30, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (11) V7-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 50, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (12) V8-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 50, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (13) V9-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5, data aug = true.
- (14) V10-S1 has hyperparameters: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, eta threshold = 0.90, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5, data aug = true.

## 5.6. 1/23/2018.

- (1) Collected results from training runs on the data\_liquid15 data. Here they are:

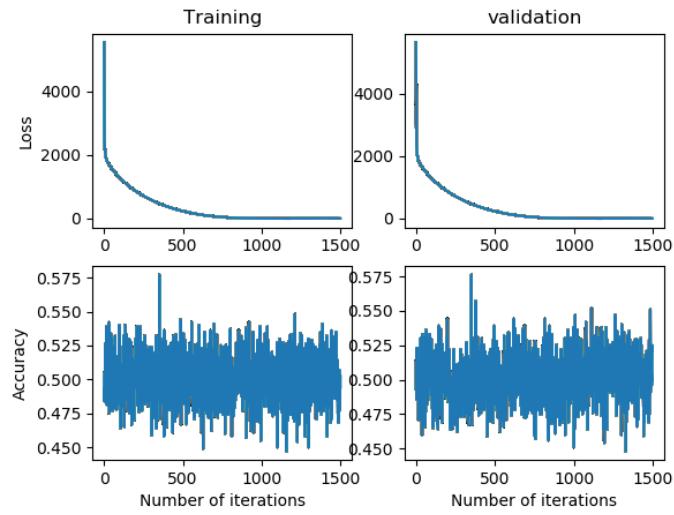


FIGURE 107. Version 1 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 10, L2 beta = 0.01, Dropout = 0.5, data aug = true

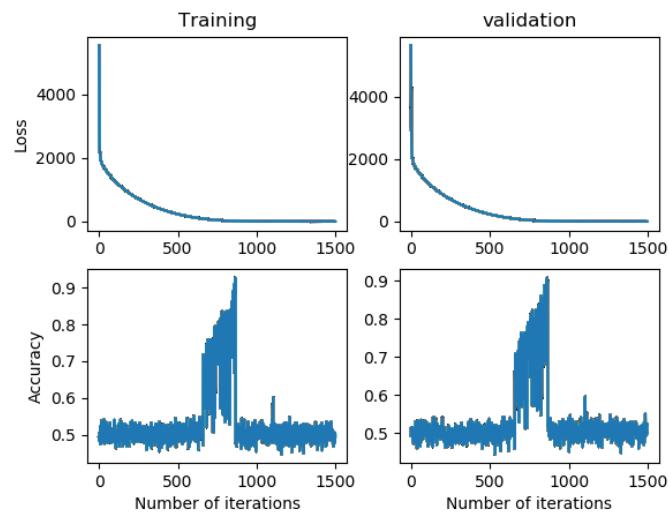


FIGURE 108. Version 2 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 10, L2 beta = 0.01, Dropout = 0.5, data aug = true

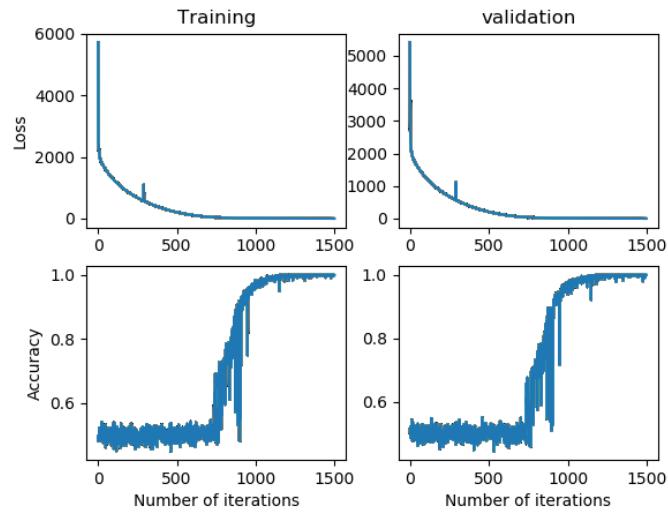


FIGURE 109. Version 3 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 20, L2 beta = 0.01, Dropout = 0.5, data aug = true

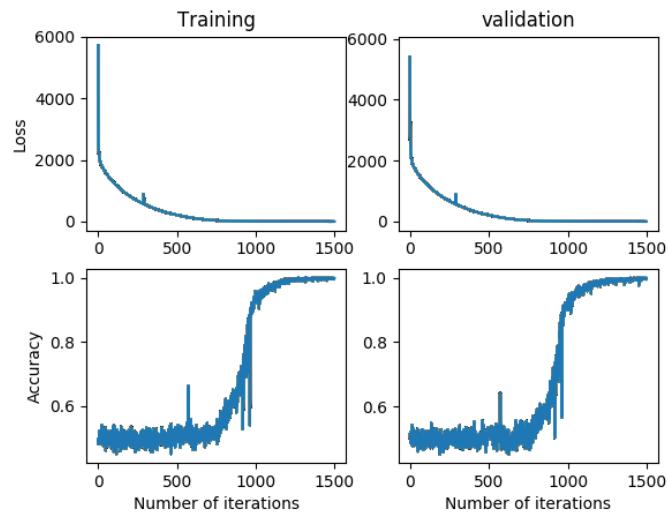


FIGURE 110. Version 4 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 20, L2 beta = 0.01, Dropout = 0.5, data aug = true

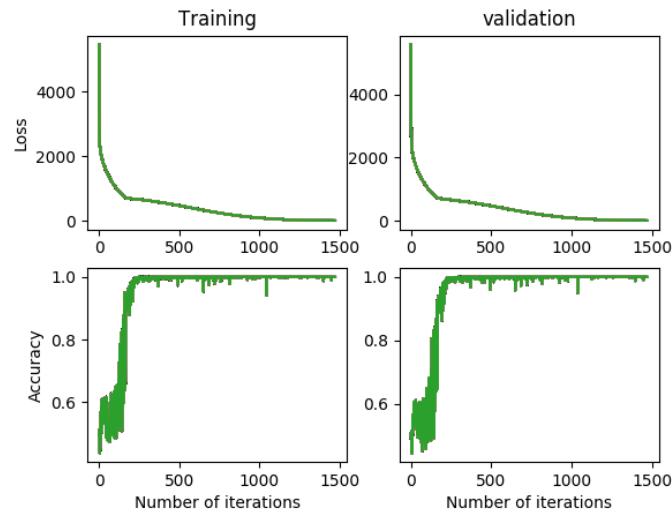


FIGURE 111. Version 5 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 30, L2 beta = 0.01, Dropout = 0.5, data aug = true

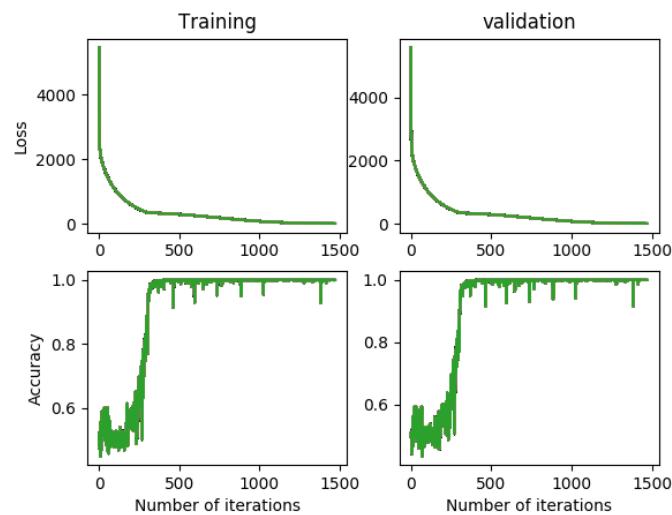


FIGURE 112. Version 6 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 30, L2 beta = 0.01, Dropout = 0.5, data aug = true

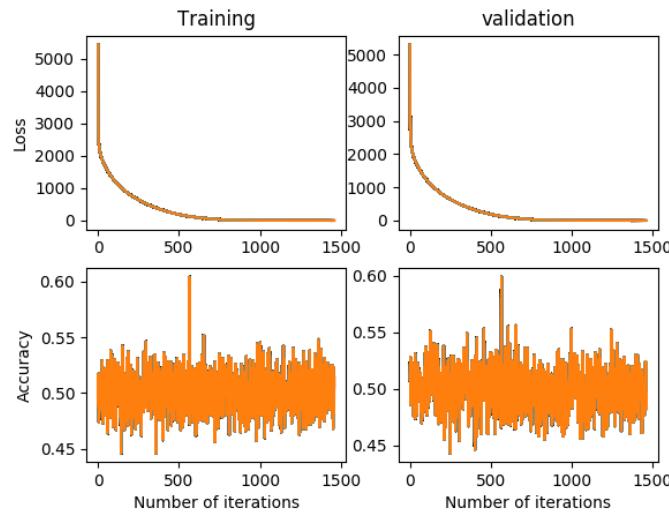


FIGURE 113. Version 7 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 50, L2 beta = 0.01, Dropout = 0.5, data aug = true

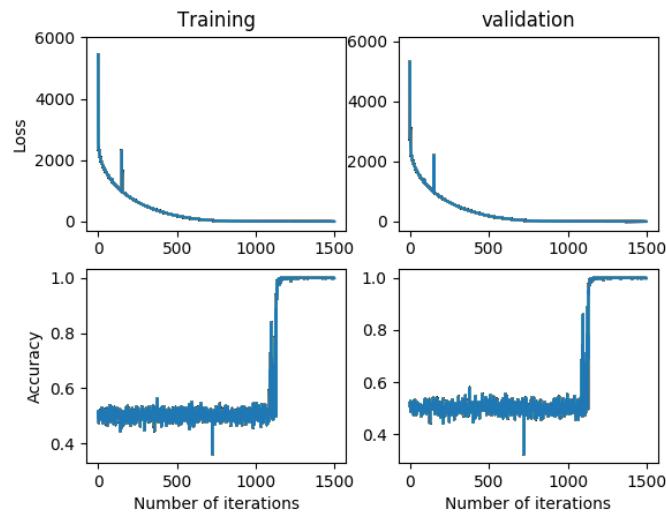


FIGURE 114. Version 8 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 50, L2 beta = 0.01, Dropout = 0.5, data aug = true

- (2) Version 4 Step 1 had a best model final validation accuracy of 99.7667% and final test accuracy of 99.9333%.
- (3) Saving this model, deleting others.
- (4) Switching metadata to data\_liquid25.
- (5) Running same set of versions as before.

### 5.7. 2/2/2018.

- (1) Collected results from training runs on the data\_liquid25 data. Here they are:

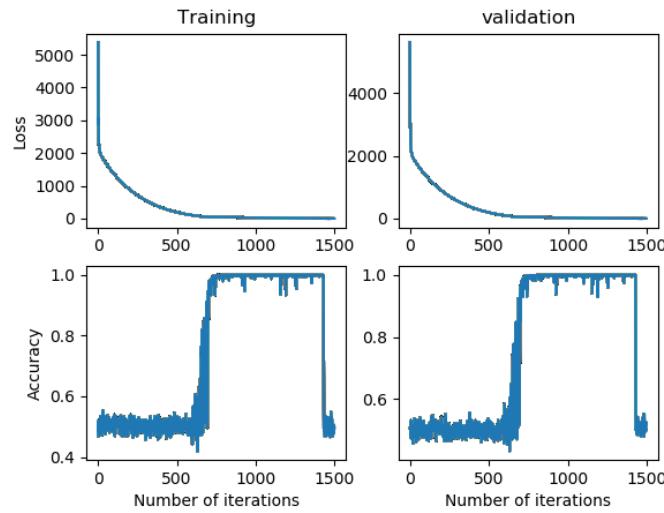


FIGURE 115. Version 1 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 10, L2 beta = 0.01, Dropout = 0.5, data aug = true

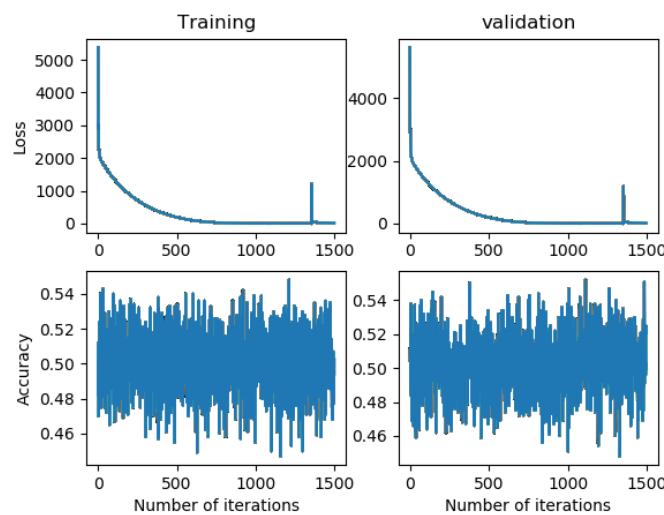


FIGURE 116. Version 2 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 10, L2 beta = 0.01, Dropout = 0.5, data aug = true

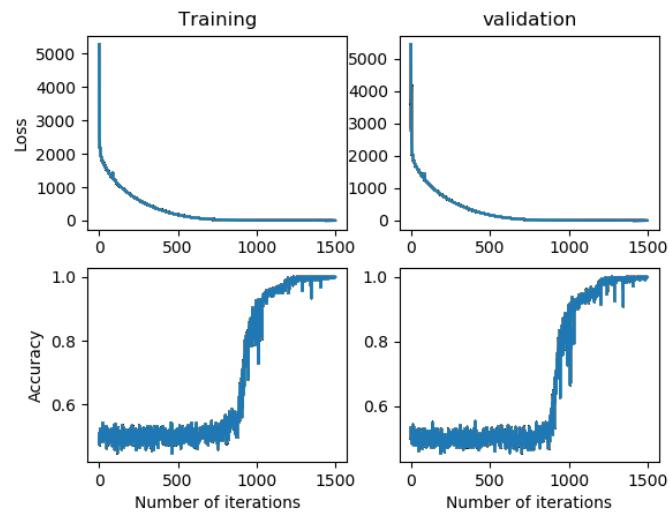


FIGURE 117. Version 3 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 20, L2 beta = 0.01, Dropout = 0.5, data aug = true

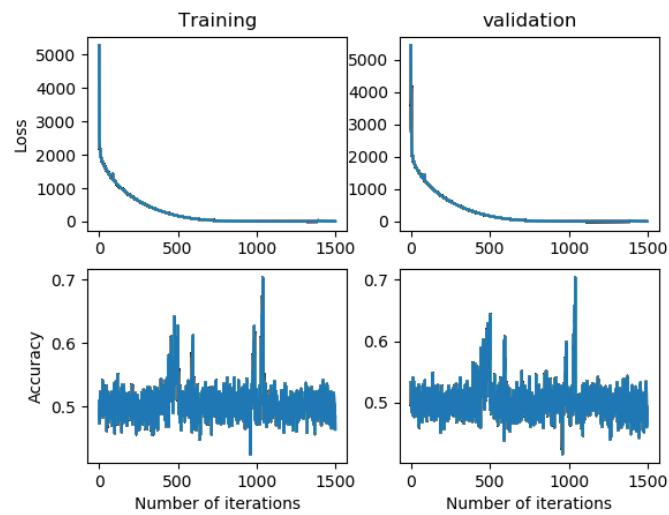


FIGURE 118. Version 4 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 20, L2 beta = 0.01, Dropout = 0.5, data aug = true

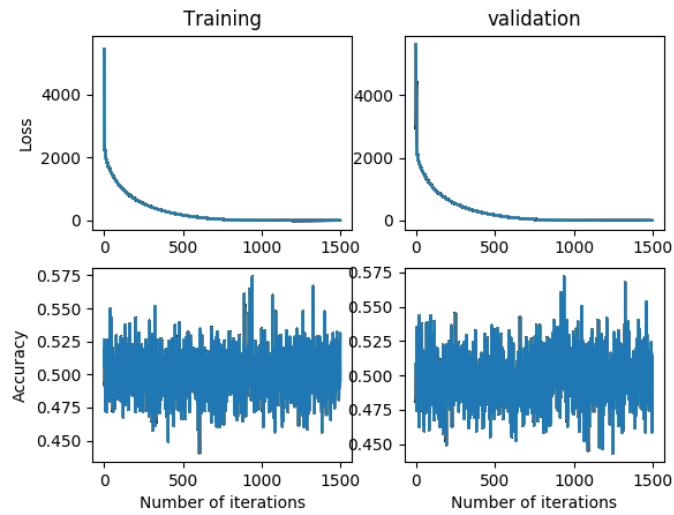


FIGURE 119. Version 5 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 30, L2 beta = 0.01, Dropout = 0.5, data aug = true

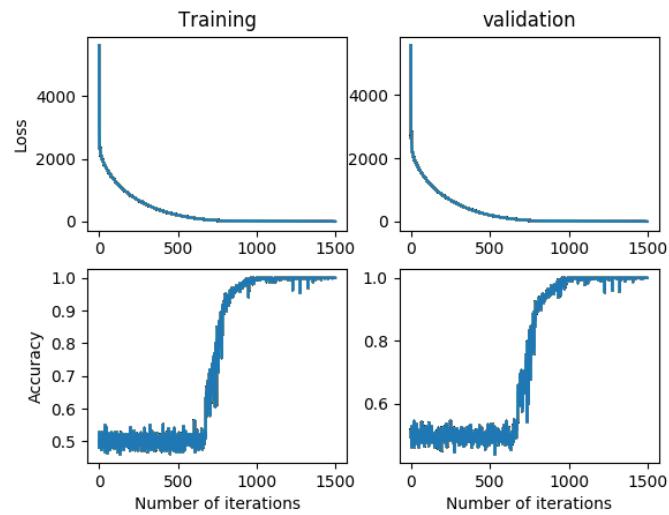


FIGURE 120. Version 6 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 30, L2 beta = 0.01, Dropout = 0.5, data aug = true

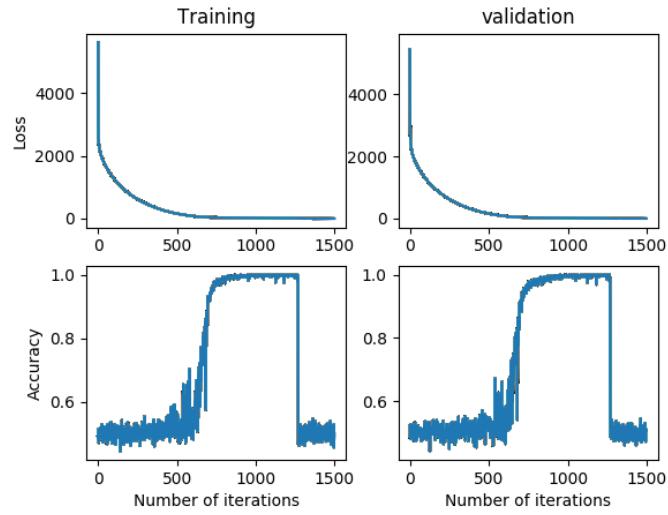


FIGURE 121. Version 7 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 50, L2 beta = 0.01, Dropout = 0.5, data aug = true

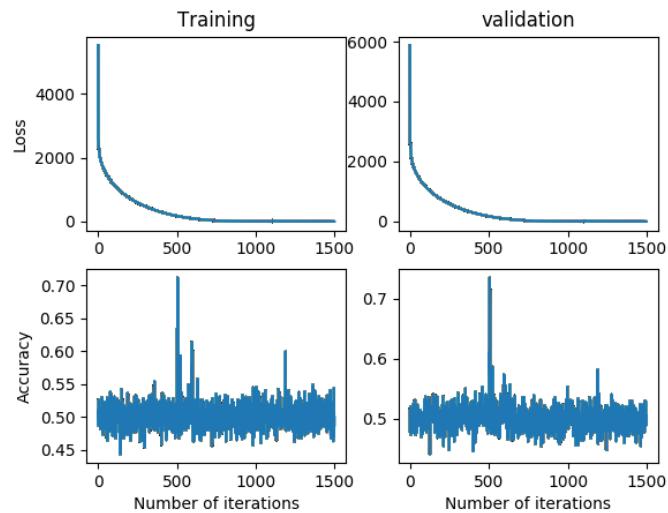


FIGURE 122. Version 8 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 50, L2 beta = 0.01, Dropout = 0.5, data aug = true

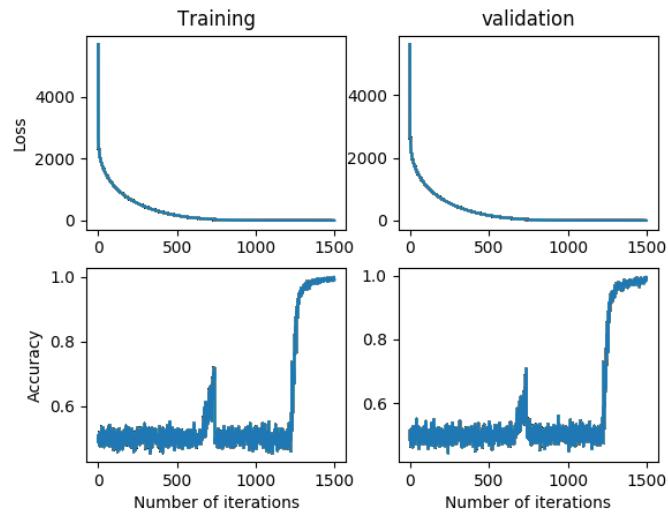


FIGURE 123. Version 8 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5, data aug = true

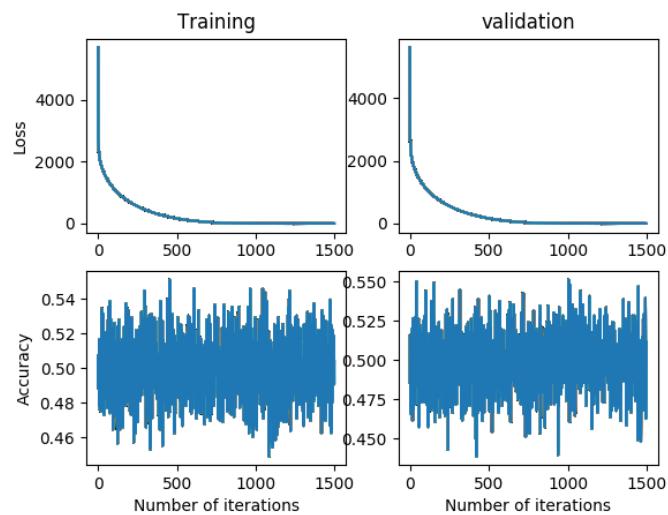


FIGURE 124. Version 8 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5, data aug = true

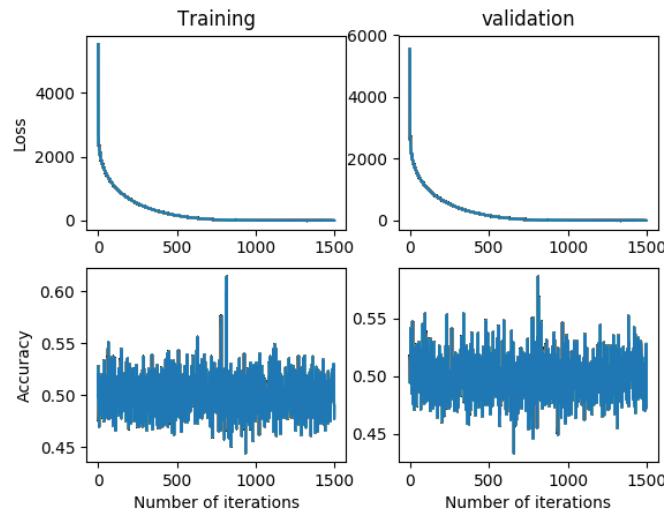


FIGURE 125. Version 8 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 100, L2 beta = 0.01, Dropout = 0.5, data aug = true

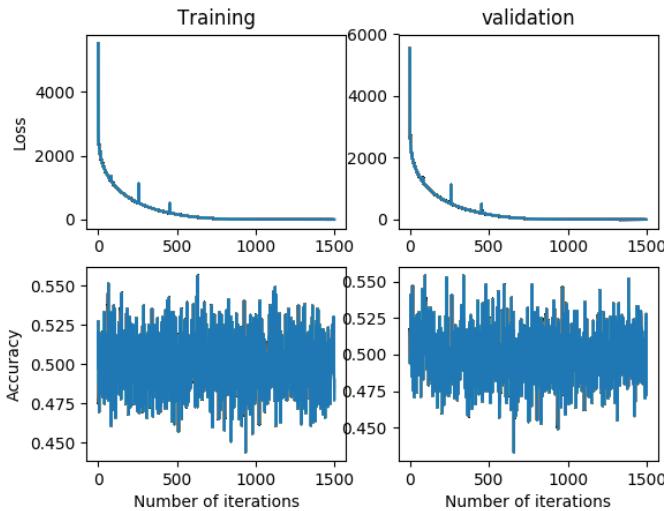


FIGURE 126. Version 8 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 100, L2 beta = 0.01, Dropout = 0.5, data aug = true

- (2) Version 6 Step 1 has a final validation accuracy of 99.7% and test accuracy of 99.6667%.
- (3) Saving Version 6 model.
- (4) Just ran the saved `data_liquid5_version7_step1_best_model.ckpt` on the metadata from `data_liquid25`. It had a final validation accuracy of 99.8557% and final test accuracy of 99.6333%. Wow! So how about for the time being, we

just keep running this model on closer and closer liquid data sets to see how well it does...

- (5) Changing metadata to be from data\_liquid35.
- (6) Running the data\_liquid5 model as above...this time gets 99.5667% validation and 99.7333% test.
- (7) Creating data now for data\_liquid45 and data\_liquid60.
- (8) Ok, now changed the metadata to be from data\_12steps. This yields a validation accuracy of 48.2667% and test accuracy of 48.5%. Interesting...so where does it break down? Let's try to find this.
- (9) Removing all other models.

### 5.8. 2/3/2018.

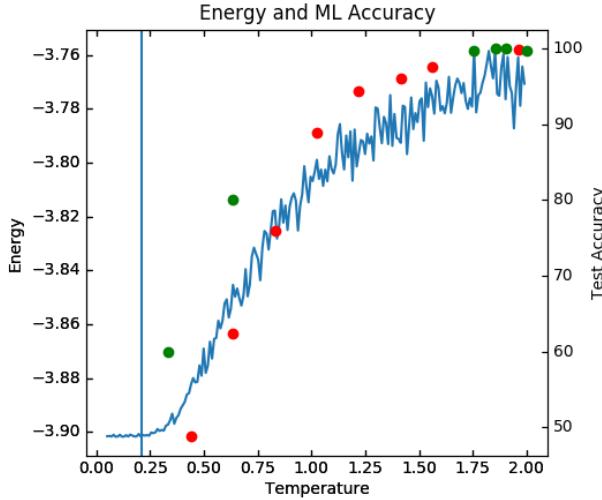
- (1) Switching metadata to be from data\_liquid45.
- (2) Running data\_liquid5\_version7\_step1\_best\_model.ckpt gives 96.4755% validation accuracy and 97.5466% test accuracy on data\_liquid45.
- (3) Switching to metadata data\_liquid60.
- (4) Running data\_liquid5\_version7\_step1\_best\_model.ckpt gives 96.4446% validation accuracy and 96.0663% test accuracy on data\_liquid60.
- (5) Switching to metadata data\_liquid80.
- (6) Running data\_liquid5\_version7\_step1\_best\_model.ckpt gives 92.8547% validation accuracy and 94.408% test accuracy on data\_liquid80.
- (7) Switching to metadata data\_liquid100.
- (8) Running data\_liquid5\_version7\_step1\_best\_model.ckpt gives 90.0206% validation accuracy and 88.8201% test accuracy on data\_liquid100.
- (9) Switching to metadata data\_liquid120.
- (10) Running data\_liquid5\_version7\_step1\_best\_model.ckpt gives 76.8569% validation accuracy and 75.9367% test accuracy on data\_liquid120.
- (11) Switching to metadata data\_liquid140.
- (12) Running data\_liquid5\_version7\_step1\_best\_model.ckpt gives 60.5473% validation accuracy and 62.3961% test accuracy on data\_liquid140.
- (13) Switching to metadata data\_liquid160.
- (14) Running data\_liquid5\_version7\_step1\_best\_model.ckpt gives 47.7124% validation accuracy and 48.7964% test accuracy on data\_liquid160.
- (15) Switching metadata to data\_liquid140 and running training again, using the same stuff as before.

### 5.9. 2/19/2018.

- (1) New idea on MPNN. Modified Josh's code, "plot\_config.py", and renamed it "coord\_config.py". This new file eliminates the plotting function, and instead outputs a text file with the raw (x, y, z) atomic coordinates of a given configuration. Added this file to GitHub.
- (2) Currently running the (x, y, z) coords program on Midway to generate these files for the endpoints of the simulation (i.e. at 20,000,000 and 100,000) for Shubhendu.

### 5.10. 2/22/2018.

(1) Here is a plot showing progress to date:



(2) What I sent to Josh and Shubhendu in email: I have a few new results. Attached to this email is a plot I put together a few minutes ago. The blue curve is the value of the inherent structure energy as a function of temperature (values on left axis) taken from one example simulation. The vertical blue line is the assumed glass transition temperature for the 2D Kob-B-Andersen glass, which is at 0.21 (from Dan's vapor-deposited paper). The dots show various machine learning accuracies I have achieved (values on the right axis). I'll explain the red dots first, then the green dots.

Red dots: I created a data set where the glass images were derived from timestep 20,000,000, or temperature = 0.05 and the liquid images were derived from timestep 500,000, or temperature = 1.95125. Temperature = 0.05 is the farthest we can go into the glass regime with our data. I then trained the CNN on this data set and achieved a test accuracy of 99.8%. I saved this model. Then, I used this model to predict whether images were glass or liquid on seven other data sets. Each of these data sets uses the SAME glass images, i.e. images from temperature = 0.05, but uses liquid images from temperatures that are increasingly close to the glass transition. The results are shown as red dots. So, the red dot at temperature = 0.635 and accuracy = 62.3961%, for example, means that this model achieves approximately 62.3961% accuracy when tested on liquid images derived from temperature = 0.635.

Green dots: These dots represent CNNs that have been trained exactly at that data point. So, the green dot at temperature = 0.33275, for example, represents the test accuracy of a CNN that has been trained on liquid images taken from that exact temperature, 0.33275, with glass images again taken from temperature = 0.05.

It looks nice, but what if we're just learning a much more complicated way of calculating the inherent structure energy?

- (3) After speaking to Josh, these are the steps we decided that we need to take, using CNNs:
- (4) Classify glass vs liquid:
  - Improve model accuracy using data across the liquid regime
  - Calculate  $T_g$  rigorously
  - Calculate the potential energy instead of the inherent structure energy
  - NEW classify between two liquids? (Added 3/6/2018)
- (5) NEW idea: identify the phase transition temperature by stepping across into the glass regime until the classification accuracy gets really low. (Added 3/8/2018)
  - Generate supercooled liquid data at the same energy as the glass
  - Generate artificial 5-fold patterns
  - Generate 3D data for classification
- (6) Classify liquid-cooled versus vapor-deposited glasses:
  - Generate data
- (7) Other tasks?
  - Classify low density vs high density water
- (8) We also have an answer as to why the CNN learning to classify liquid versus glass, irrespective of what it is learning (i.e. potential energy, inherent energy, geometry, or anything else). From the viewpoint of the CNN, we know nothing about the material that is being given to the CNN other than the geometric configuration of particles. We don't know whether interactions are pairwise or more complicated, we don't know how the material operates, other than that we have been able to train the CNN on labelled materials, where the labelling could be based on some macroscopic rather than microscopic property. And yet, the CNN is still able to take just this raw structure information, no dynamics, and still tell the difference between a glass and a liquid. So it is able to classify these materials with high accuracy based on nothing but structural atomic configurations. This is interesting.
- (9) So, based on these above steps, I will first start by running more training to improve green dot accuracy.

### 5.11. 2/27/2018.

- (1) Logging results now that should have been logged a while ago. This is the training results for data\_liquid140.

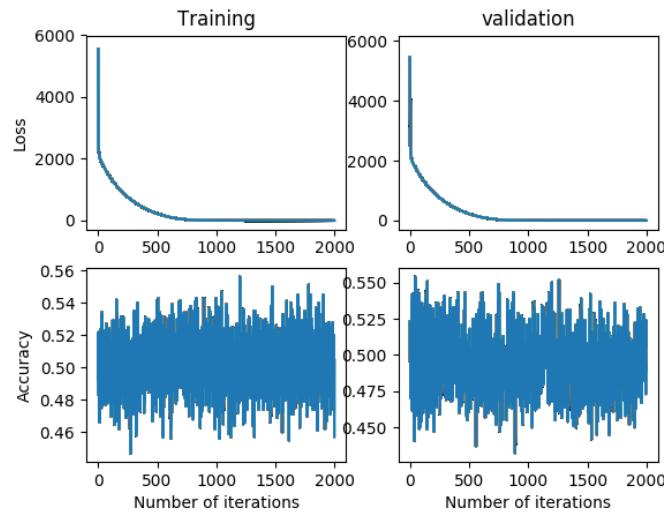


FIGURE 127. Version 1 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 10, L2 beta = 0.01, Dropout = 0.5, data aug = true

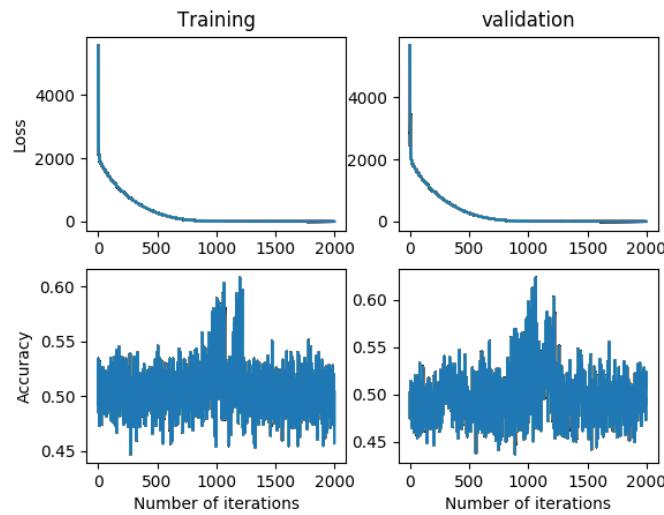


FIGURE 128. Version 2 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 10, L2 beta = 0.01, Dropout = 0.5, data aug = true

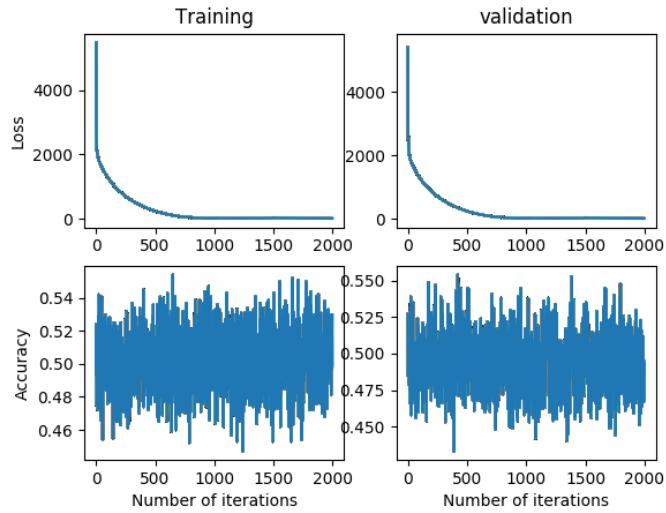


FIGURE 129. Version 3 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 20, L2 beta = 0.01, Dropout = 0.5, data aug = true

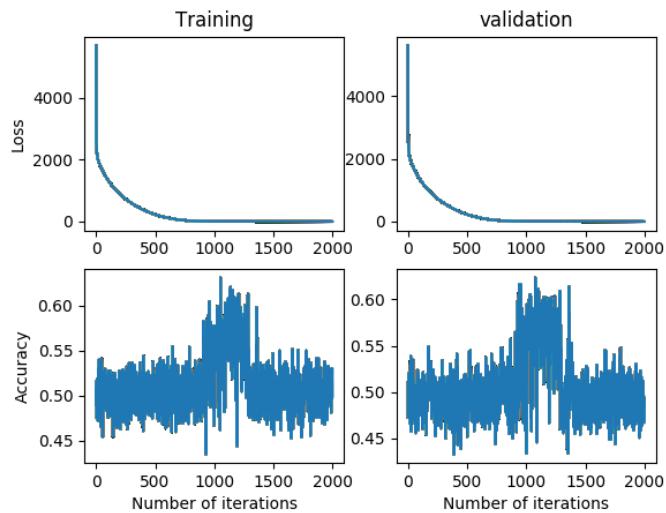


FIGURE 130. Version 4 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 20, L2 beta = 0.01, Dropout = 0.5, data aug = true

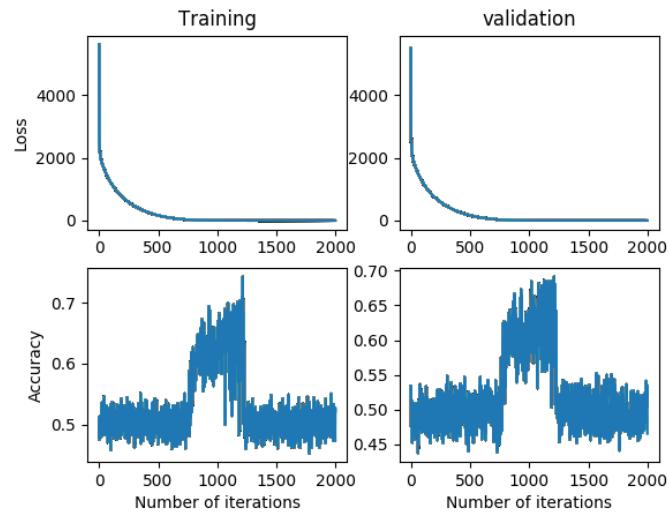


FIGURE 131. Version 5 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 30, L2 beta = 0.01, Dropout = 0.5, data aug = true

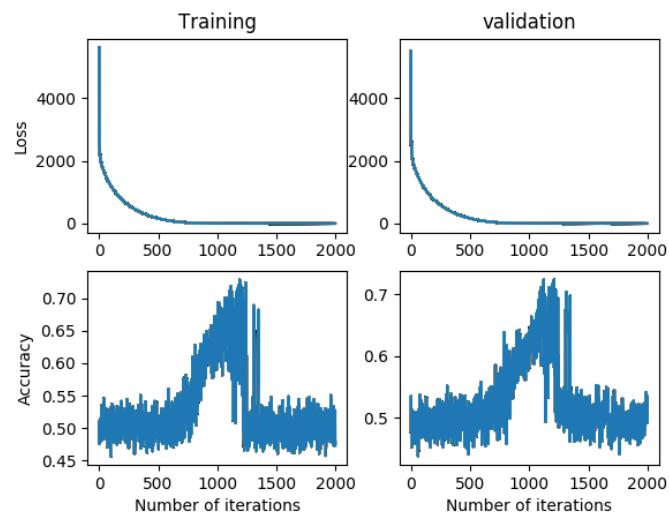


FIGURE 132. Version 6 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 30, L2 beta = 0.01, Dropout = 0.5, data aug = true

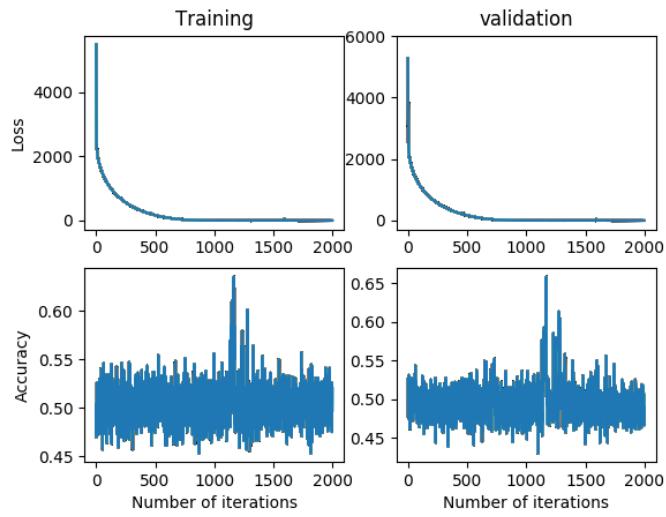


FIGURE 133. Version 7 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 50, L2 beta = 0.01, Dropout = 0.5, data aug = true

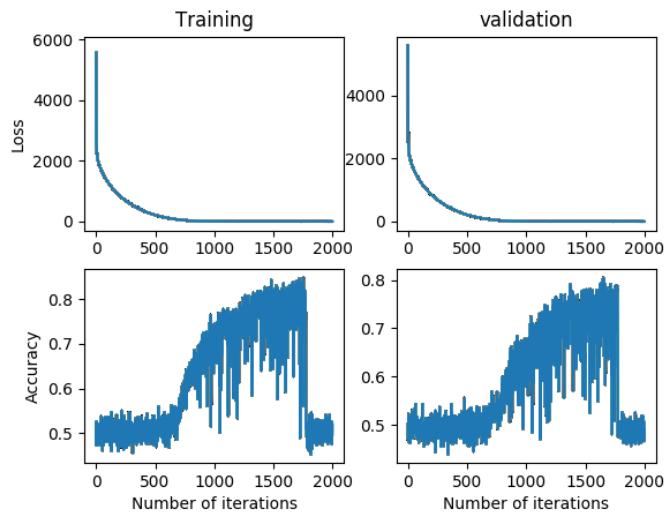


FIGURE 134. Version 8 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 50, L2 beta = 0.01, Dropout = 0.5, data aug = true

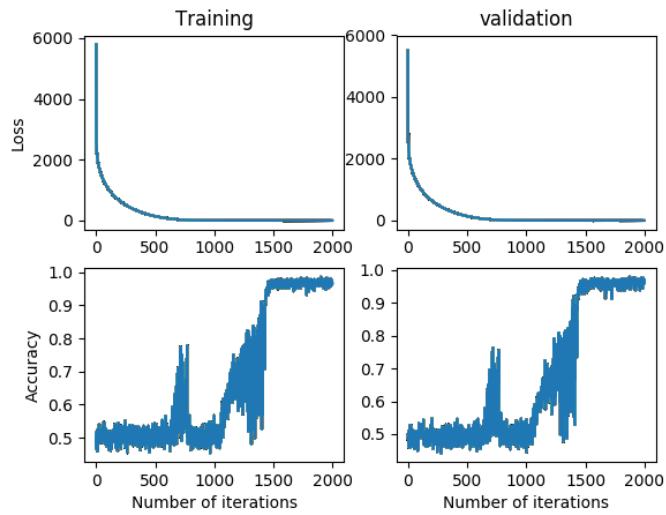


FIGURE 135. Version 9 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5, data aug = true

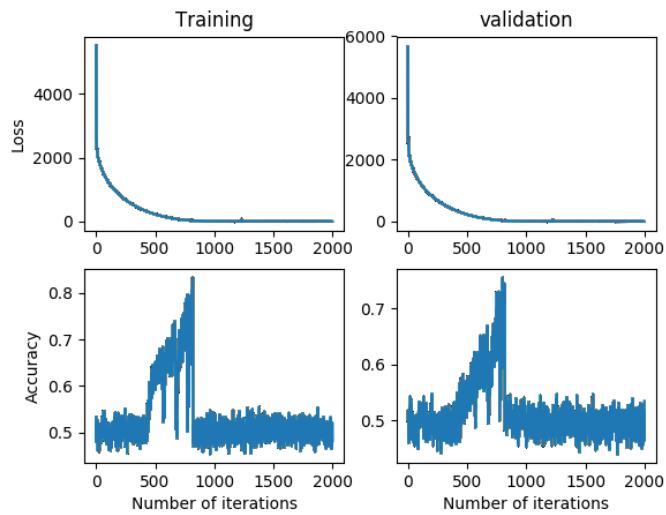


FIGURE 136. Version 10 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5, data aug = true

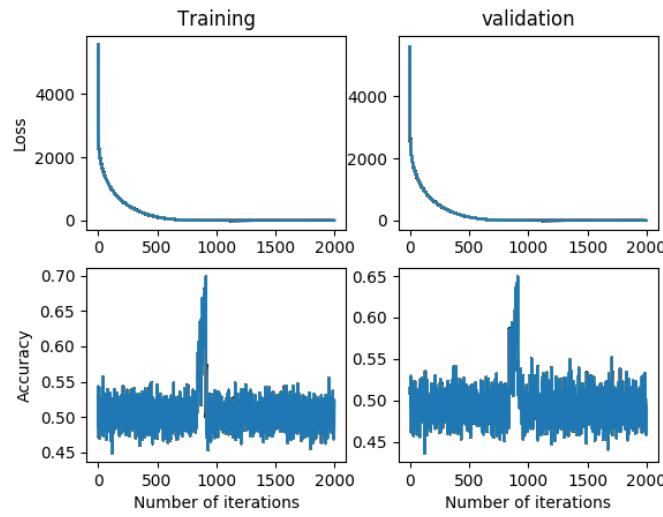


FIGURE 137. Version 11 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 100, L2 beta = 0.01, Dropout = 0.5, data aug = true

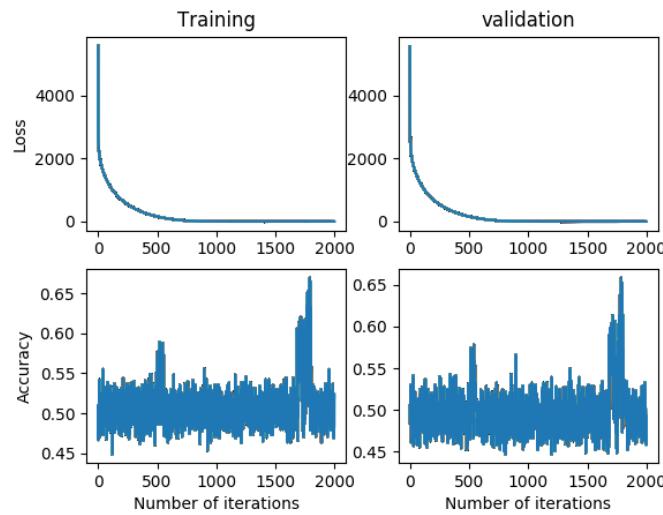


FIGURE 138. Version 12 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 100, L2 beta = 0.01, Dropout = 0.5, data aug = true

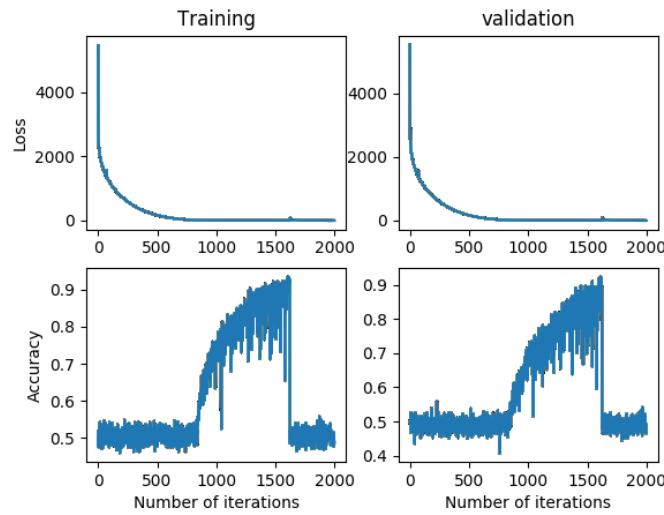


FIGURE 139. Version 13 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 130, L2 beta = 0.01, Dropout = 0.5, data aug = true

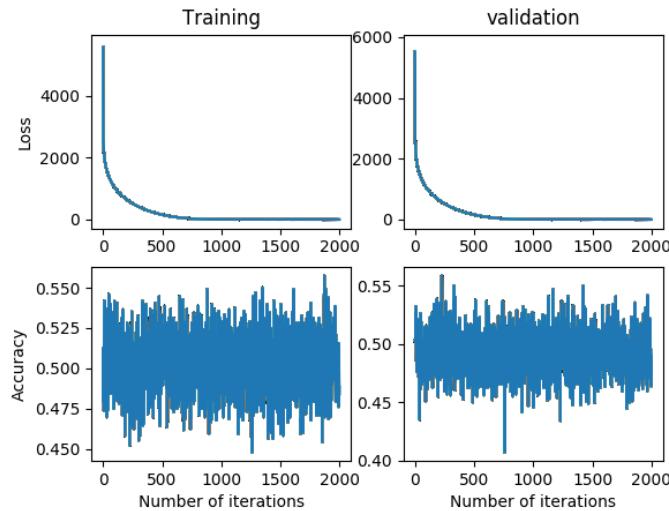


FIGURE 140. Version 14 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 130, L2 beta = 0.01, Dropout = 0.5, data aug = true

- (2) Version 9 Step 1 has a final validation accuracy of 96.2847% and test accuracy of 96.3194%.
- (3) Saving Version 9 Step 1 model. Deleting other information.
- (4) Now I want to find the true average glass transition temperature for the simulation data that we have. What I will do is construct an average cooling.dat plot, i.e. a plot that averages the data from all 10,000 cooling.dat files that we have.

Then, I will estimate by eye (or using some slightly more quantitative method for sake of argument) what the true average glass transition temperature is. We anticipate that it is higher than 0.21.

### 5.12. 3/6/2018.

- (1) Wrote a program called plot\_average\_cooling data, which I moved to the 2017/N2e7 directory. It collects the data from all 10,000 cooling.dat files and averages them across each temperature, so that we obtain an average cooling curve that is representative of the simulations:

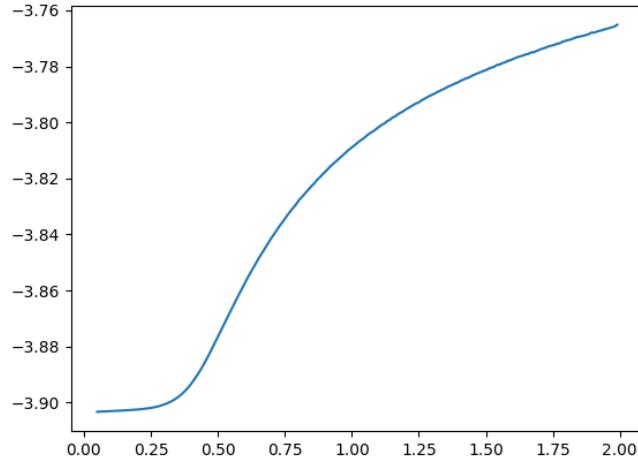


FIGURE 141. Averaged cooling.dat data

- (2) Estimated the kink in inherent structure energy by finding the intersection of two lines, using points (0.05, -3.9033) and (0.245, -3.902) for the glass data and points (0.44, -3.8876) and (0.557, -3.8654) for the liquid regime. The result is a glass transition at around 0.37. Removing the data\_12steps result, as this is actually stepping over the new glass transition, here is a new plot of our results:

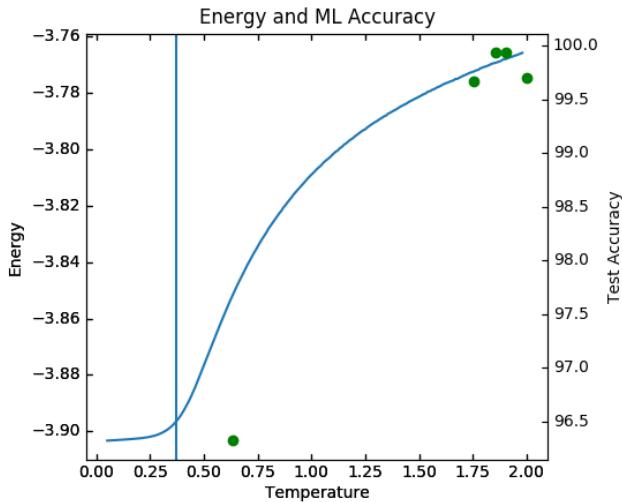


FIGURE 142. Updated results plot

- (3) Switching to data from data\_liquid155, which corresponds to  $T = 0.479$ .
- (4) Now running the same versions as before.
- (5) PROBLEM: How is this any different from distinguishing between two liquids at different energies/temperatures? SO we should try to classify between two liquids!!! Added to the to do list above.

#### 5.13. 3/7/2018.

- (1) Read sections of papers (now on Box) on MPNN, meeting with Shubhendu tomorrow.
- (2) Still running training on 155.
- (3) Looked into simulating new liquids that will be at the same inherent structure energy as the glass, but ran into issues with the batch scripts as written - emailed Josh for assistance.

#### 5.14. 3/8/2017.

- (1) New idea: identify the phase transition temperature by stepping across into the glass regime until the classification accuracy gets really low.
- (2) Reading <https://arxiv.org/pdf/1802.09876.pdf>. Suggests that networks with one layer are not only sufficient, but more interpretable.
- (3) Analysis of activation of neurons. Previously, for Ising model, only 1 layer with 3 neurons needed. Plotting weight matrices of each neuron. Look at activations as a function of five-fold symmetry??

#### 5.15. 3/9/2018.

- (1) We will be implementing MPNN using: [https://github.com/priba/nmp\\_qc](https://github.com/priba/nmp_qc)
- (2) First instruction is to install requirements using pip install -r requirements.txt
- (3) On Midway 2, /home/swansonk1 directory, error in installing torch: "PyTorch does not currently provide packages for PyPI"

- (4) For the time being, I will try this on my laptop instead. Actually, my desktop, which has cuda installed.
- (5) Again, pip install -r requirements.txt failed, so for torch, i.e. PyTorch, I used the command conda install pytorch torchvision -c pytorch from pytorch.org.
- (6) Successfully installed all packages in the requirements.txt file.
- (7) Downloaded the qm9 dataset.
- (8) Following Sec. 6 on <https://arxiv.org/pdf/1704.01212.pdf>, the input feature for our particles will be a one-hot encoding of either particle A or particle B. The adjacency matrix will be a raw distance feature matrix, where the entries are one-dimensional and indicate the euclidean distance between the pair of atoms.
- (9) Tried running python main.py, but saying no module rdkit.
- (10) So, went to <http://www.rdkit.org/docs/Install.html> and conda installed. Now I can create an environment as described. However, torch no longer works inside the environment. Need to investigate how environments work, then.
- (11) Opened the rdkit environment. Then re-installed PyTorch as above.
- (12) No module named networkx. So, pip install networkx within the rdkit environment. Same thing for joblib and tensorboard\_logger.
- (13) Tried running main.py. Got error, 'Graph' method has no attribute 'nodes\_iter'. Found this solution: <https://stackoverflow.com/questions/33734836/graph-object-has-no-attribute-nodes-iter-in-networkx-module-python>. Trying it now by replacing with nodes() in datasets/utils.py, specifically on line 28 in the qm9\_nodes function.
- (14) Also removing edges\_iter on line 54. Issue seems to be that the new version of networkx has mismatches.
- (15) Having some more issues with GPU compatibility. PyTorch says my GPU is too old. Also, there are more inscrutable errors. So, I am going to uninstall networkx and install the version that should be compatible with this model, networkx version1.11. Used pip install networkx==1.11.
- (16) New error with .expand in line 174 of m\_mpnn. Stack suggests that this has to do with bug in new version of PyTorch: <https://github.com/pytorch/pytorch/issues/2491>. So, like networkx, I am going to find the versin of PyTorch that was latest when the last commit for this program was made.
- (17) Found that the most recent version of PyTorch prior to the last commit on MPNN May 29, 2017, was the version v0.1.12: <https://github.com/pytorch/pytorch/releases>, on May 2, 2017. So, I am installing this version using conda install pytorch=0.1.12 -c soumith. This lets me download without cuda: <http://pytorch.org/previous-versions/>. Hopefully this will get rid of the GPU problem? If not, we'll turn to Midway.
- (18) Ok, this works! I was just able to run, in the demos folder, demo\_qm9\_mpnn.py. Now, I am going to go through this and try to understand what's going on. If necessary, I will add comments to this code file.
- (19) The data is transformed into special data objects via the Qm9 class in the datasets folder. Each object of this class has four main attributes: the graph, the nodes, the edges, and the targets. The graph is the adjacency matrix that describes connections between atoms in the molecule. It is real symmetric and is a numpy matrix. The nodes are given as a list of features for each particle. The edges

are described by a dictionary, where (1, 2): [distance, one-hot encoding of bonding] is the format. Finally, the target properties are given as a list as well.

- (20) We would probably choose our adjacency graph to take into account only a certain number of nearest neighbors as "bonded." Then, we would just need to make sure that the graph is **connected**. That is the only requirement of the scheme, which can then later be optimized.

#### 5.16. 3/13/2018.

- (1) Modified the coords\_config.py file so that it outputs atom type, x-coordinate, and y-coordinate instead of x, y, z.
- (2) Produced a sample file from the file N2e7/0 at time 100,000
- (3) Wrote a script adjacency\_matrix.py to compute the adjacency graph and relevant information for MPNN input. It works generally, but the loops appear to be inefficient. Next time, I will have to look into using matrices, as well as ensuring that the matrix is **connected**.

#### 5.17. 3/14/2018.

- (1) Added some functionality to adjacency\_matrix. Found functions that could be used for matrix manipulations, namely euclidean\_distances from sklearn and arr.argsort() for numpy arrays. Also found that you can load txt into numpy array using np.loadtxt().

#### 5.18. 3/15/2018.

- (1) Created graph\_reader.py, which returns graph information from an input glass/liquid data file in less than 7 seconds. Used matrix operations rather than only loops to increase computational efficiency.
- (2) Tomorrow, I need to go to RCC to check how to make the MPNN code work on Midway, and I also need to implement a check for connected graphs.
- (3) Downloading results from data\_liquid155.

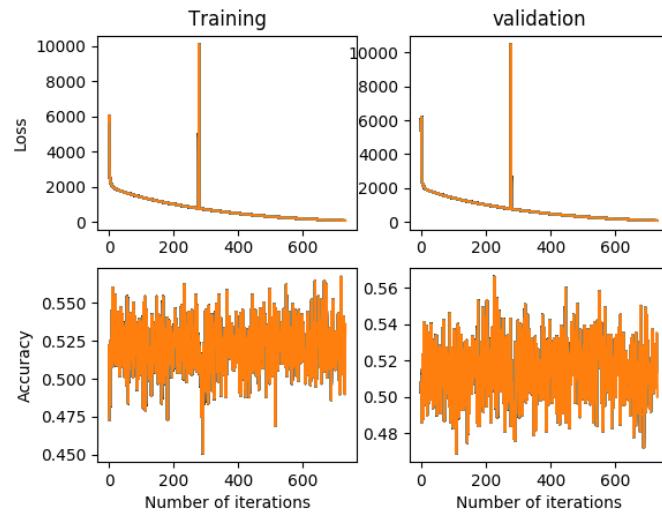


FIGURE 143. Version 1 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 10, L2 beta = 0.01, Dropout = 0.5, data aug = true

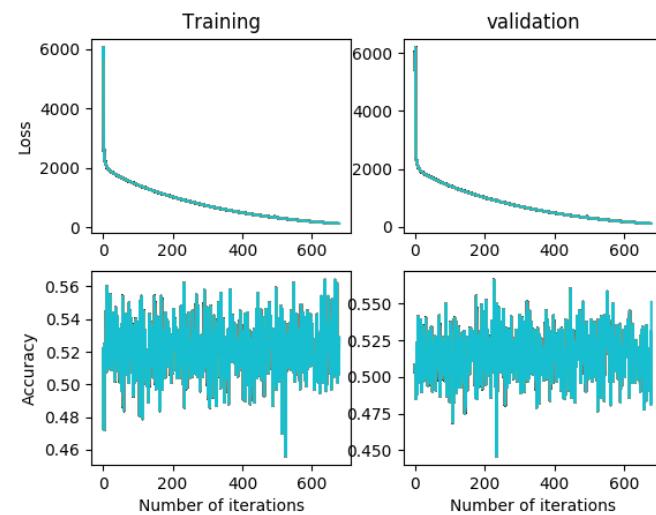


FIGURE 144. Version 2 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 10, L2 beta = 0.01, Dropout = 0.5, data aug = true

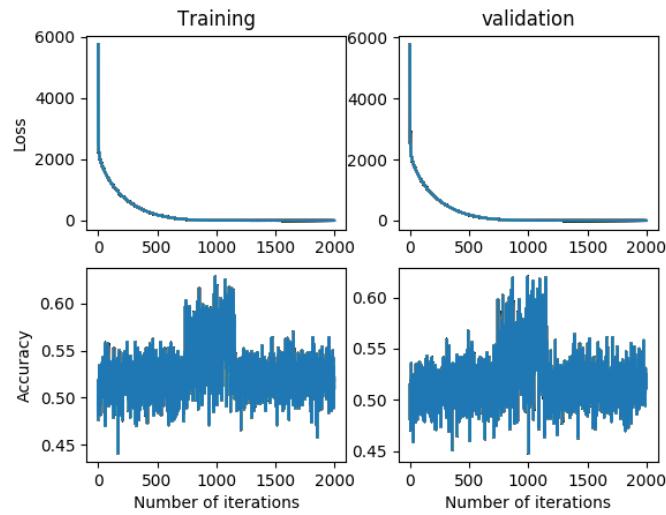


FIGURE 145. Version 3 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 20, L2 beta = 0.01, Dropout = 0.5, data aug = true

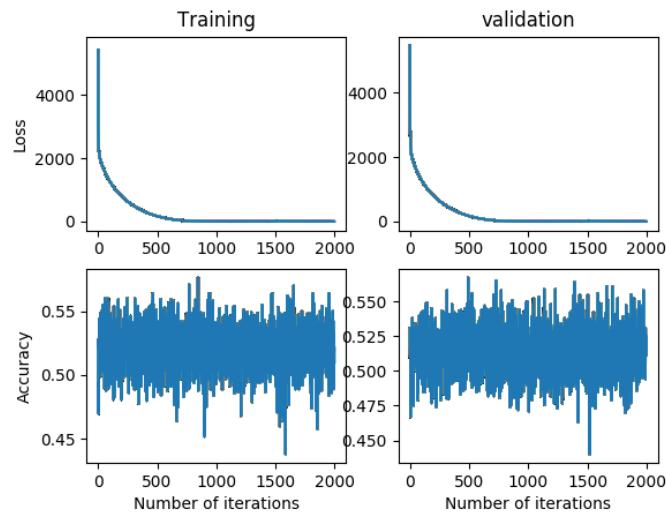


FIGURE 146. Version 4 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 20, L2 beta = 0.01, Dropout = 0.5, data aug = true

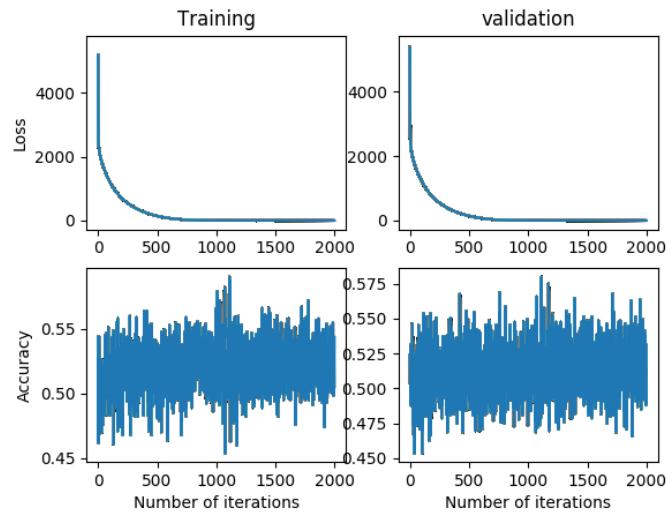


FIGURE 147. Version 5 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 30, L2 beta = 0.01, Dropout = 0.5, data aug = true

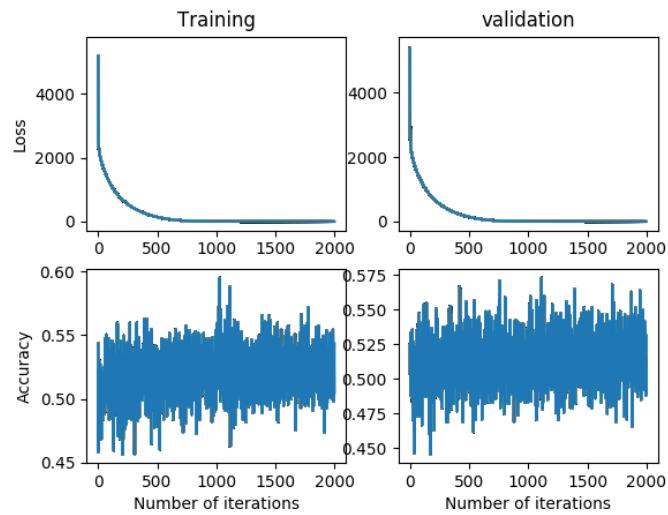


FIGURE 148. Version 6 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 30, L2 beta = 0.01, Dropout = 0.5, data aug = true

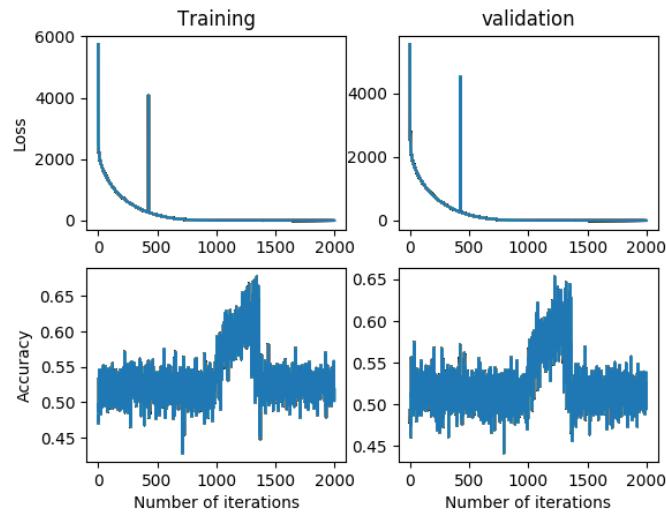


FIGURE 149. Version 7 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 50, L2 beta = 0.01, Dropout = 0.5, data aug = true

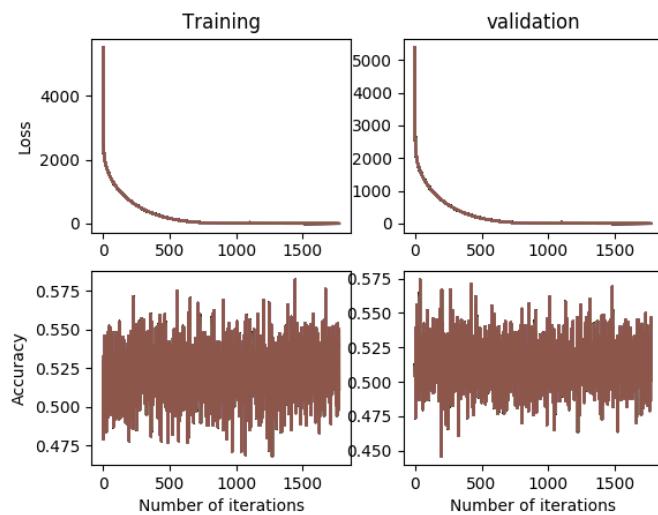


FIGURE 150. Version 8 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 50, L2 beta = 0.01, Dropout = 0.5, data aug = true

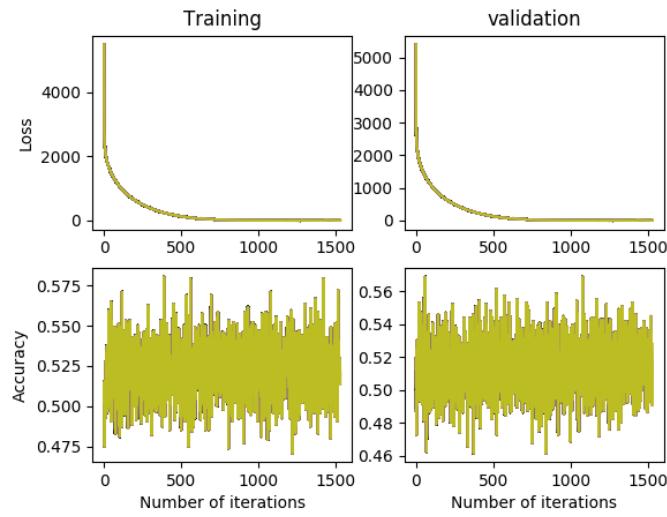


FIGURE 151. Version 9 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5, data aug = true

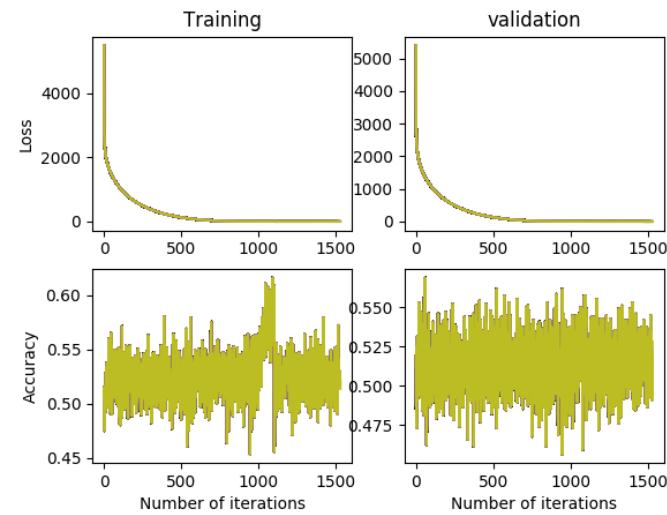


FIGURE 152. Version 10 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 80, L2 beta = 0.01, Dropout = 0.5, data aug = true

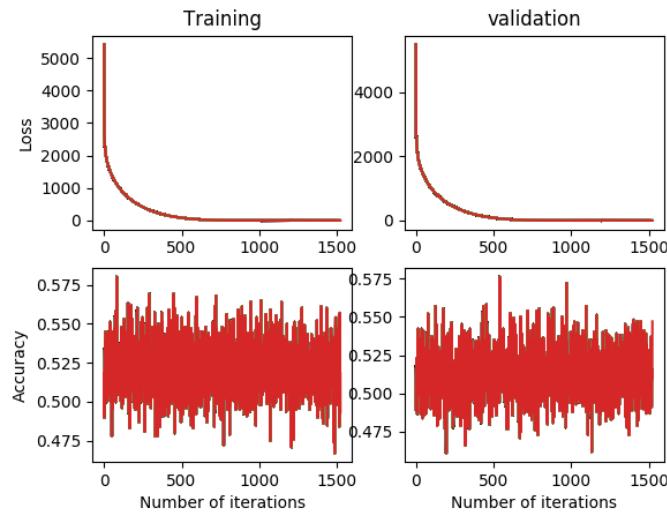


FIGURE 153. Version 11 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 100, L2 beta = 0.01, Dropout = 0.5, data aug = true

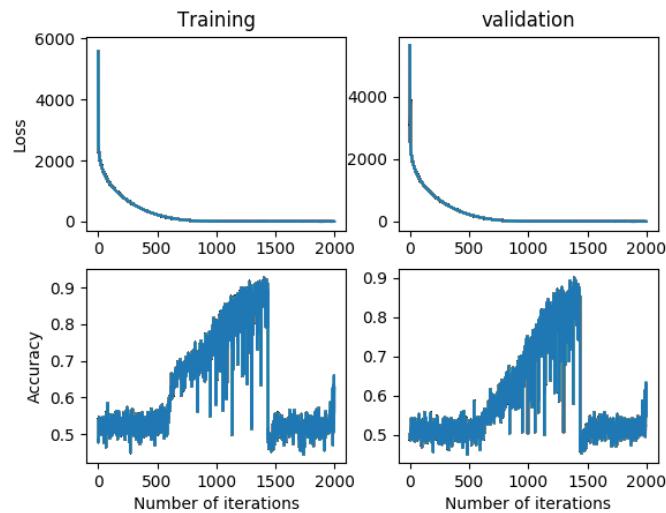
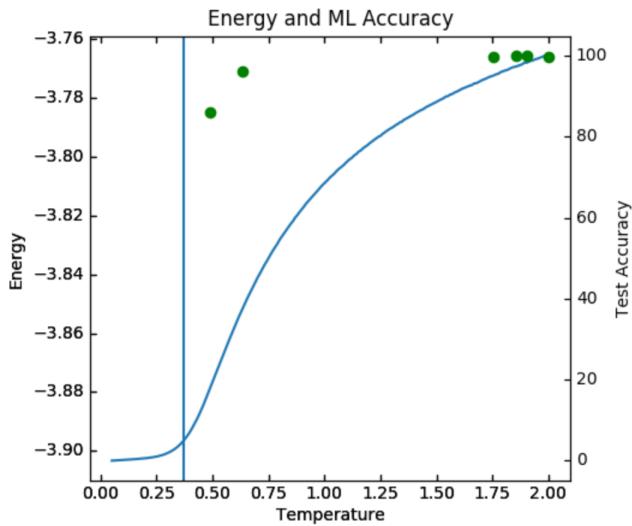


FIGURE 154. Version 12 step 1: Initial Learning Rate = 1e-3, Final Learning Rate = 1e-4, Eta threshold = 0.90, Batch Size = 100, L2 beta = 0.01, Dropout = 0.5, data aug = true

- (4) Version 12 Step 1 has a final validation accuracy of 86.9408% and test accuracy of 86.1111%. These could have probably been improved using a lower learning rate threshold.
- (5) Saving Version 12 Step 1 model. Deleting other information.
- (6) Updated accuracies:



## REFERENCES

- [R] D. Reid, I. Lyubimov, M. D. Ediger, J. J. de Pablo, *Age and structure of a model vapour-deposited glass*, Nature Communications, 2016.
- [QM] J. Gilmer, S. Schoenholz, P. Riley, O. Vinyals, and G. Dahl, *Neural Message Passing for Quantum Chemistry*, www.arxiv.org, 2017.
- [GC] S. Kearnes, K. McCloskey, M. Berndl, V. Pande, and P. Riley, *Molecular Graph Convolutions: Moving Beyond Fingerprints*, www.arxiv.org, 2017.
- [P] Q. Wei, R. G. Melko, and Jeff Z. Y. Chen, *Identifying polymer states by machine learning*, www.arxiv.org, 2017.

*E-mail address:* swansonk1@uchicago.edu

INSTITUTE FOR MOLECULAR ENGINEERING, UNIVERSITY OF CHICAGO, 5640 S ELLIS AVE,  
CHICAGO, IL 60637