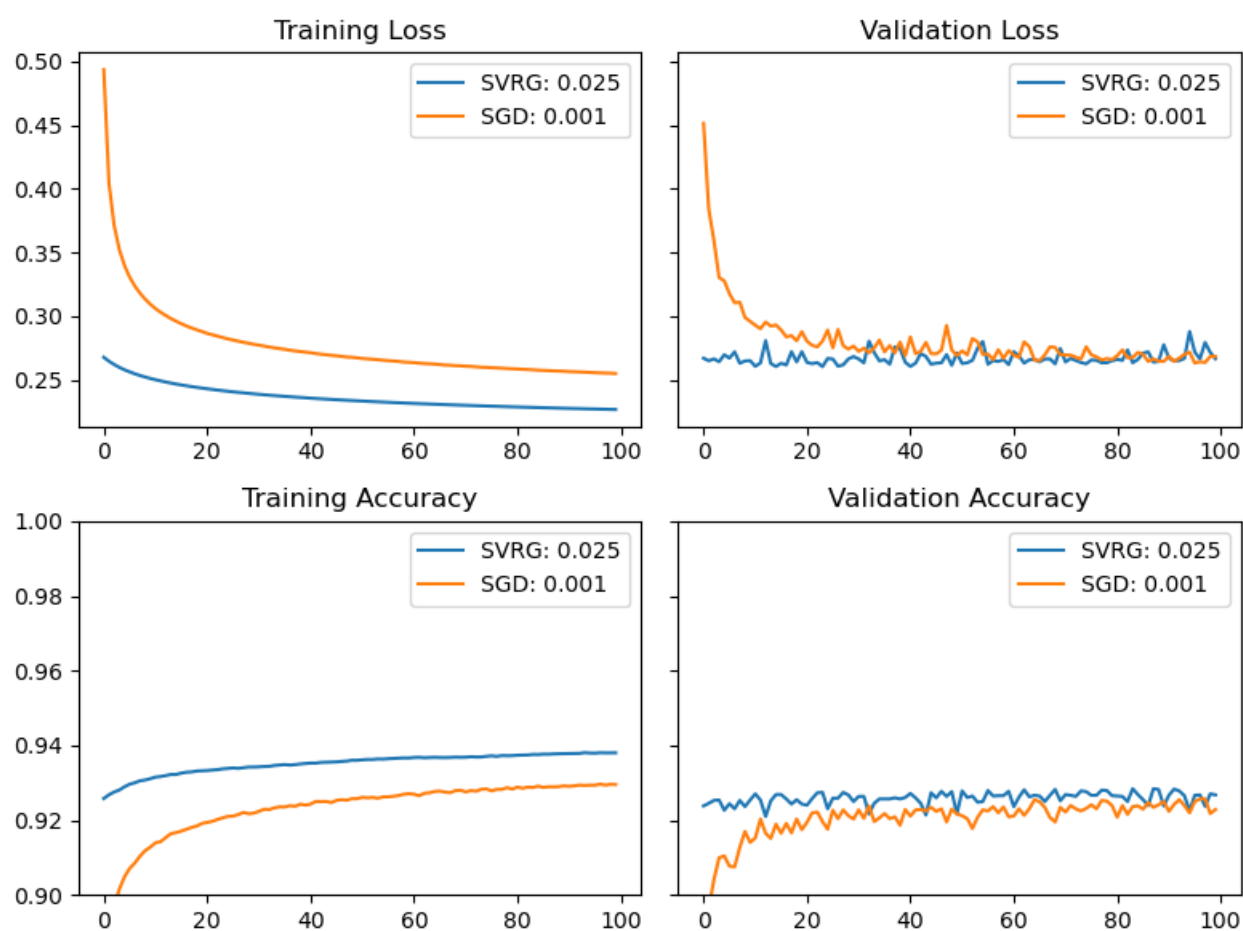


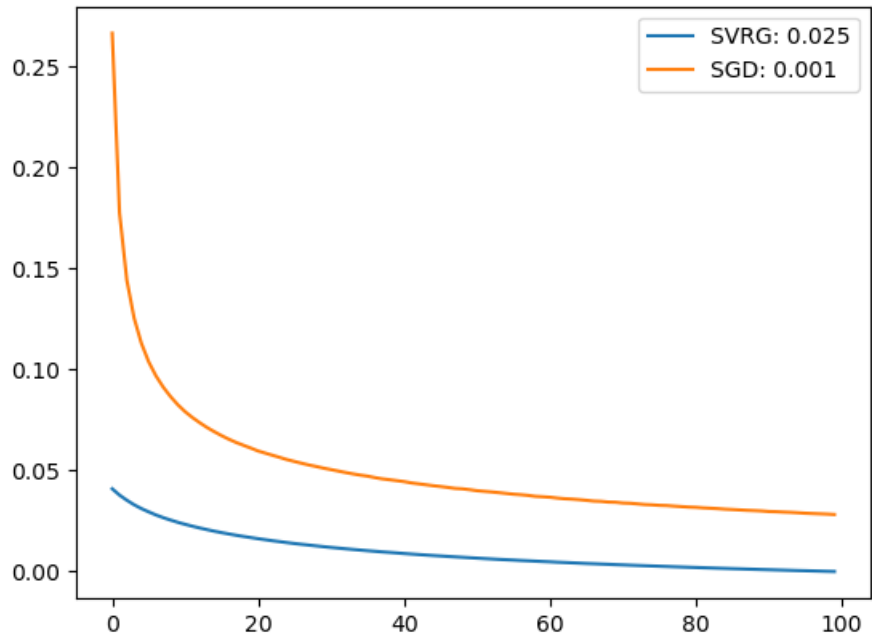
Hw1 Report

(a.)

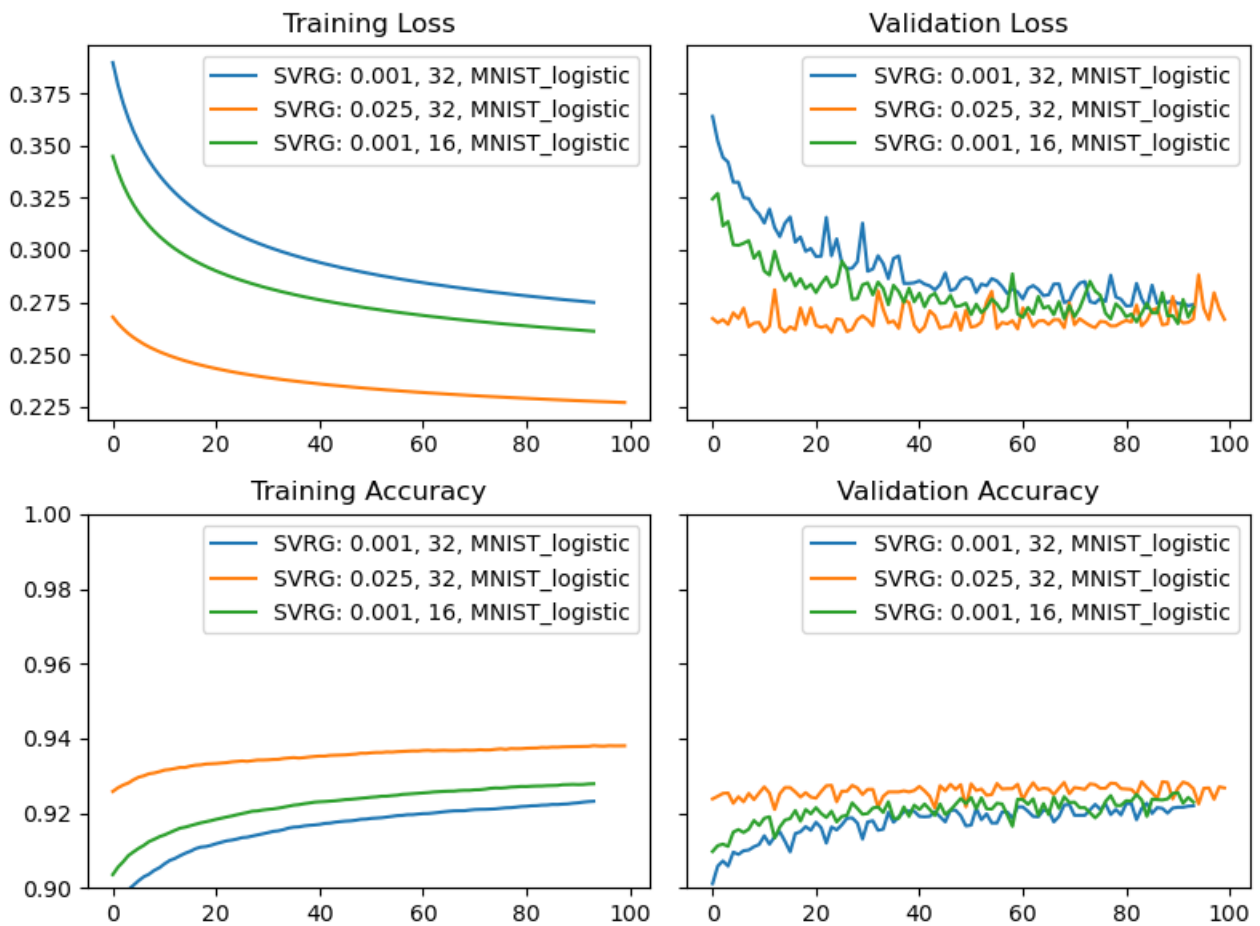
First I try to use iteration ratio with 1:1 to draw the result of SGD and SVRG graph, but I found that SVRG will convergence to fast rather than the result on the paper, so I try to make it like 3:1 in my graph. Outstandingly, it seems like it should be less than this ratio, but I think it is close to the answer in the paper. All graph below use batch size as 32.



And here is the loss residual graph about SGD and SVRG

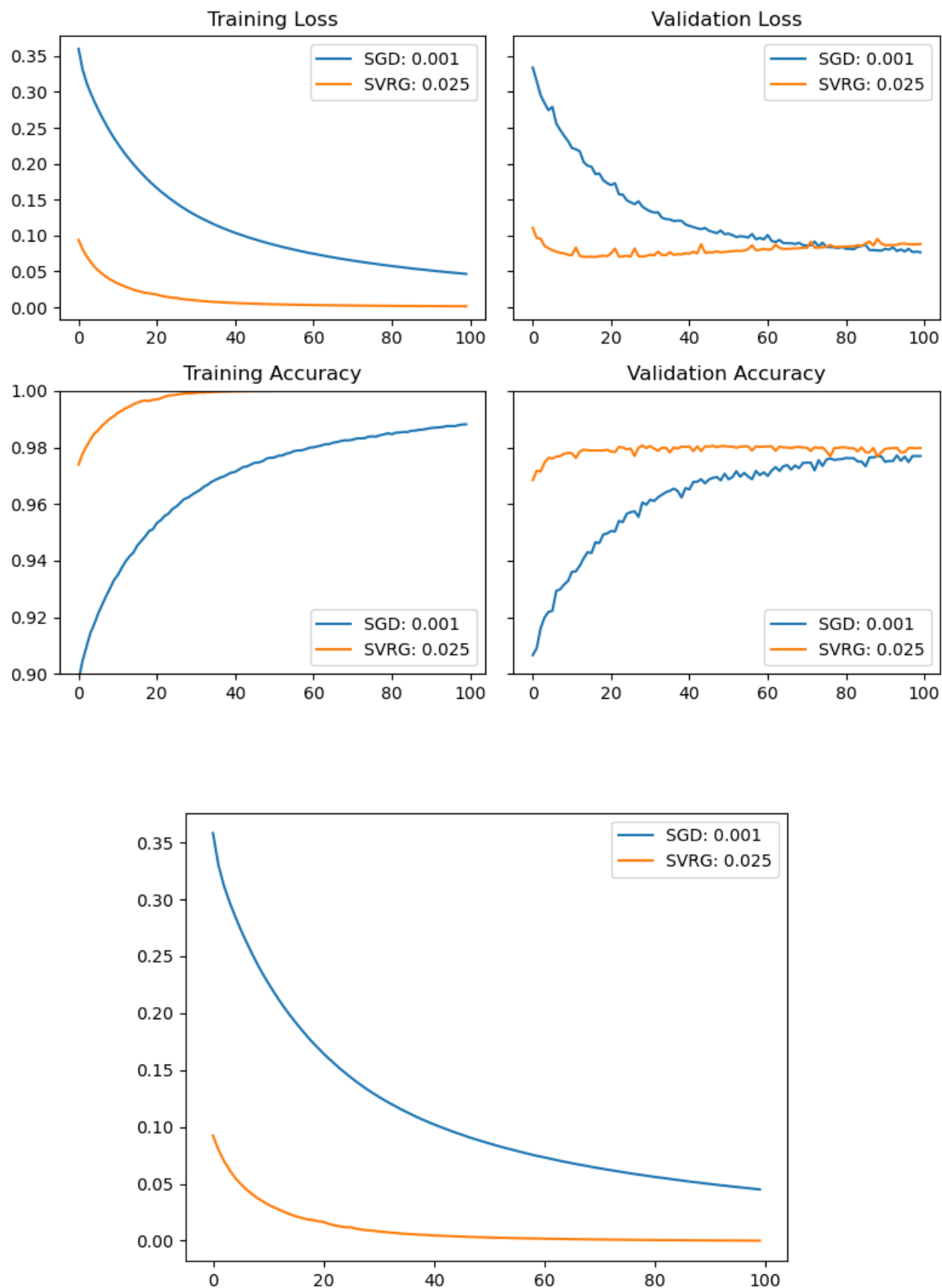


Graph about the SVRG lr and batch size on logistic_model.



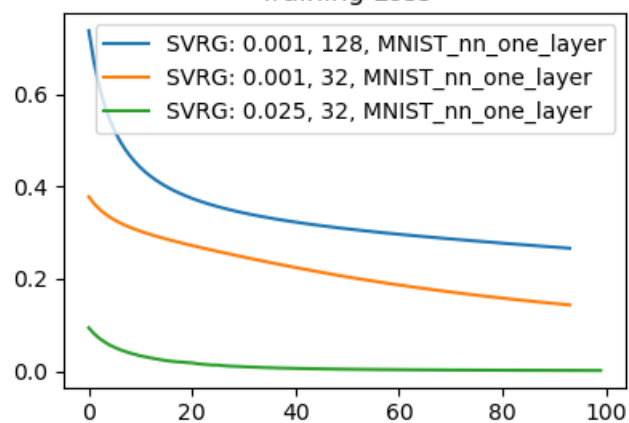
(b.)

The result of nn_linear_model with the ratio 3:1

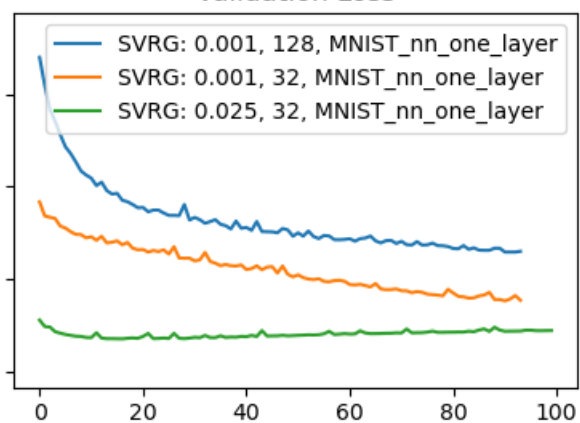


Graph about the SVRG lr and batch size on nn_model.

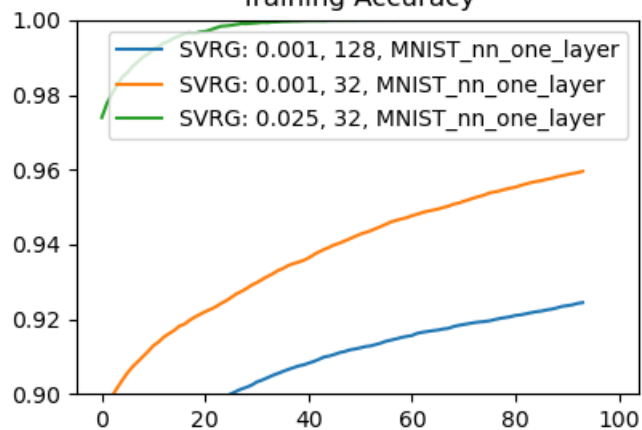
Training Loss



Validation Loss



Training Accuracy



Validation Accuracy

