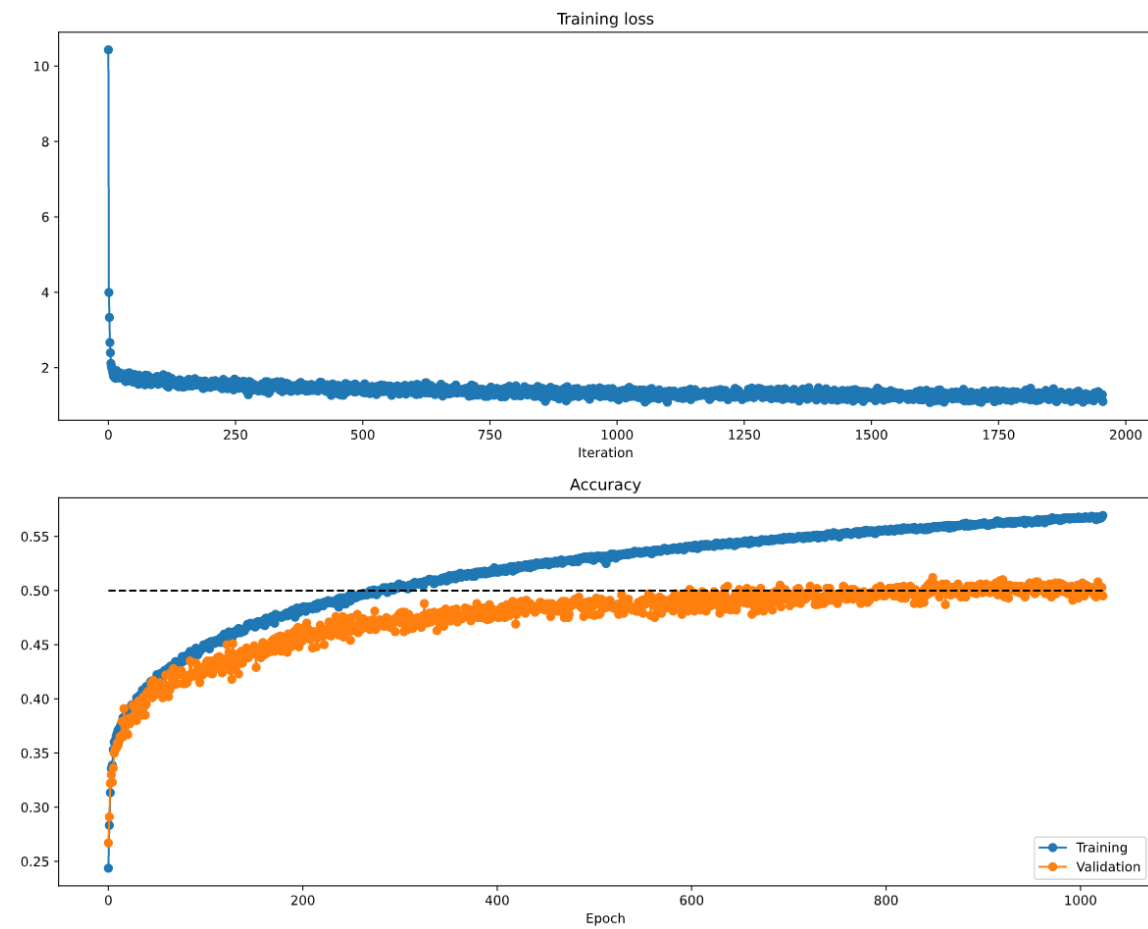


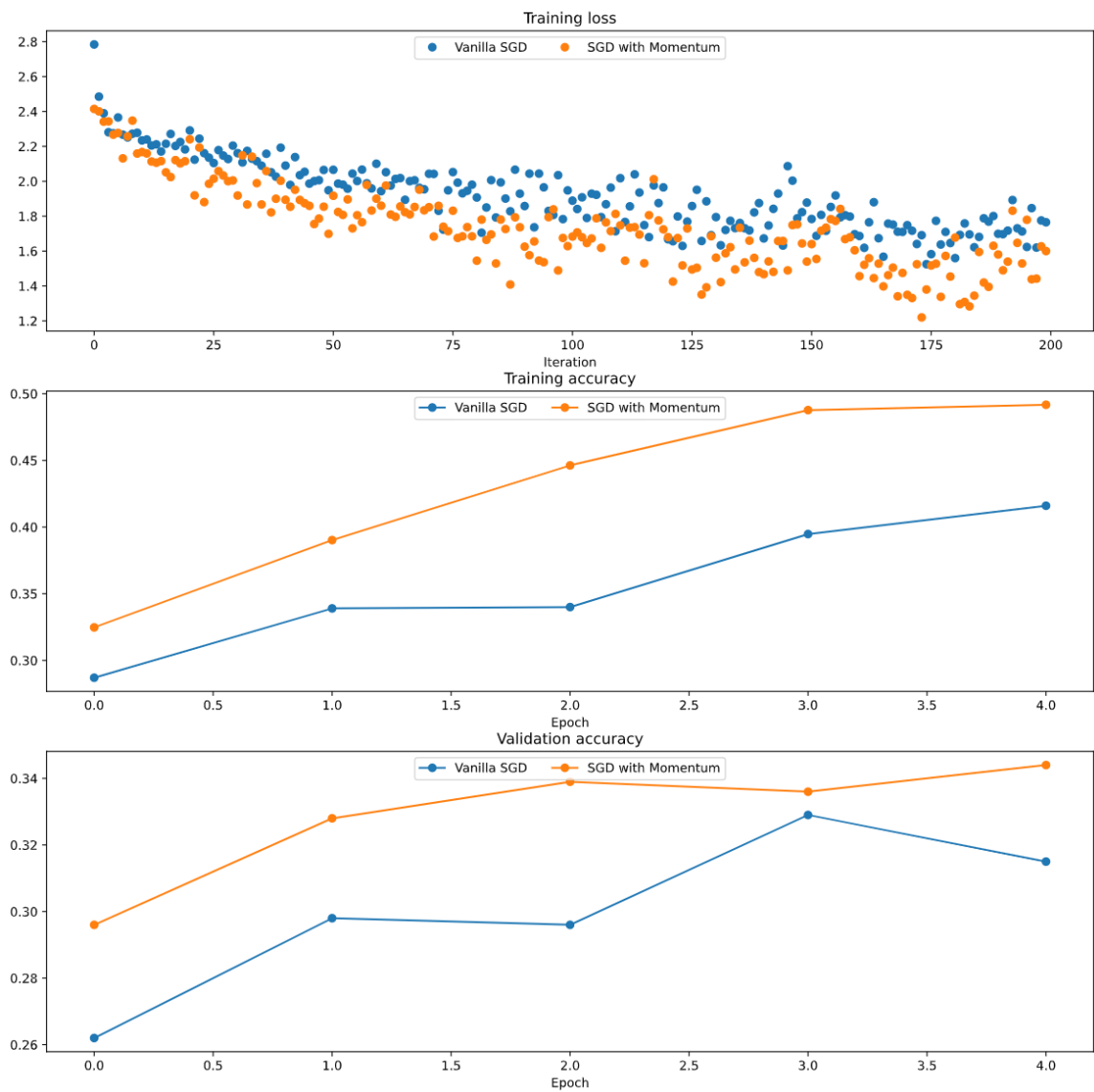
Problem 1 Solution

Lei Ding

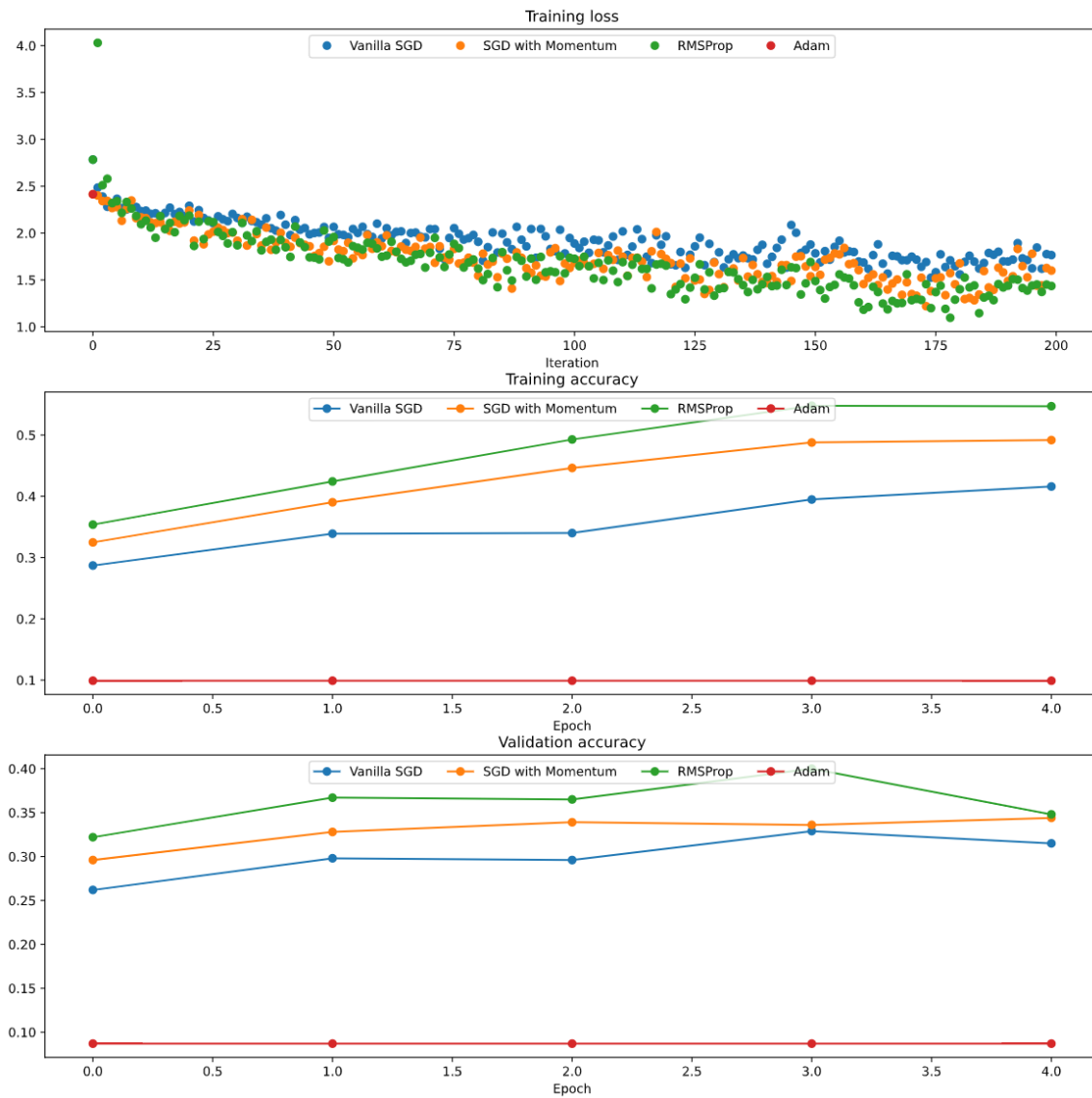
1. Training loss / accuracy for simple neural network22



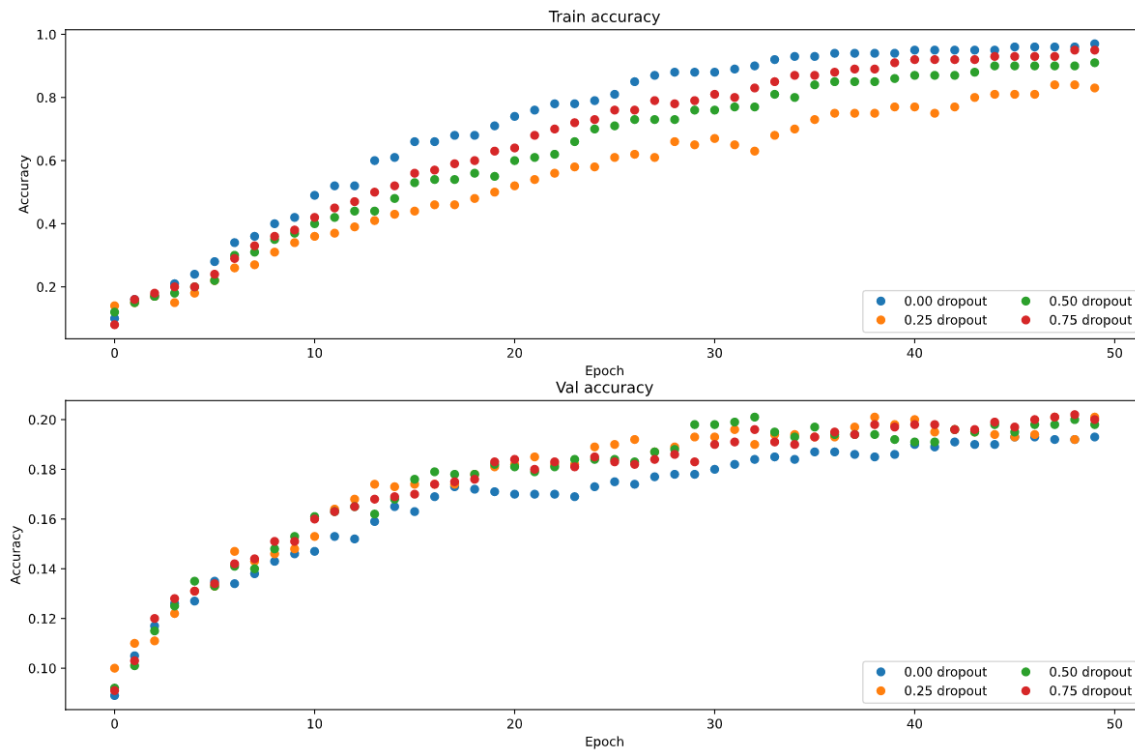
2. Vanilla SGD vs SGD with Momentum



3. Optimizer Comparison



4. Network with Dropout



5. Dropout question:

Higher dropout rate has lower training accuracy, but higher validation accuracy. It is possibly due to randomly less adjustable parameters during training. The model would be less fit for the training data because the model size is smaller. However, this prevents overfitting as some parameters won't be adjusted towards certain data. However, as training goes on for more epochs, the parameters seem to be smoothed out and the difference between different dropout rates gets smaller.

6. Activation functions

