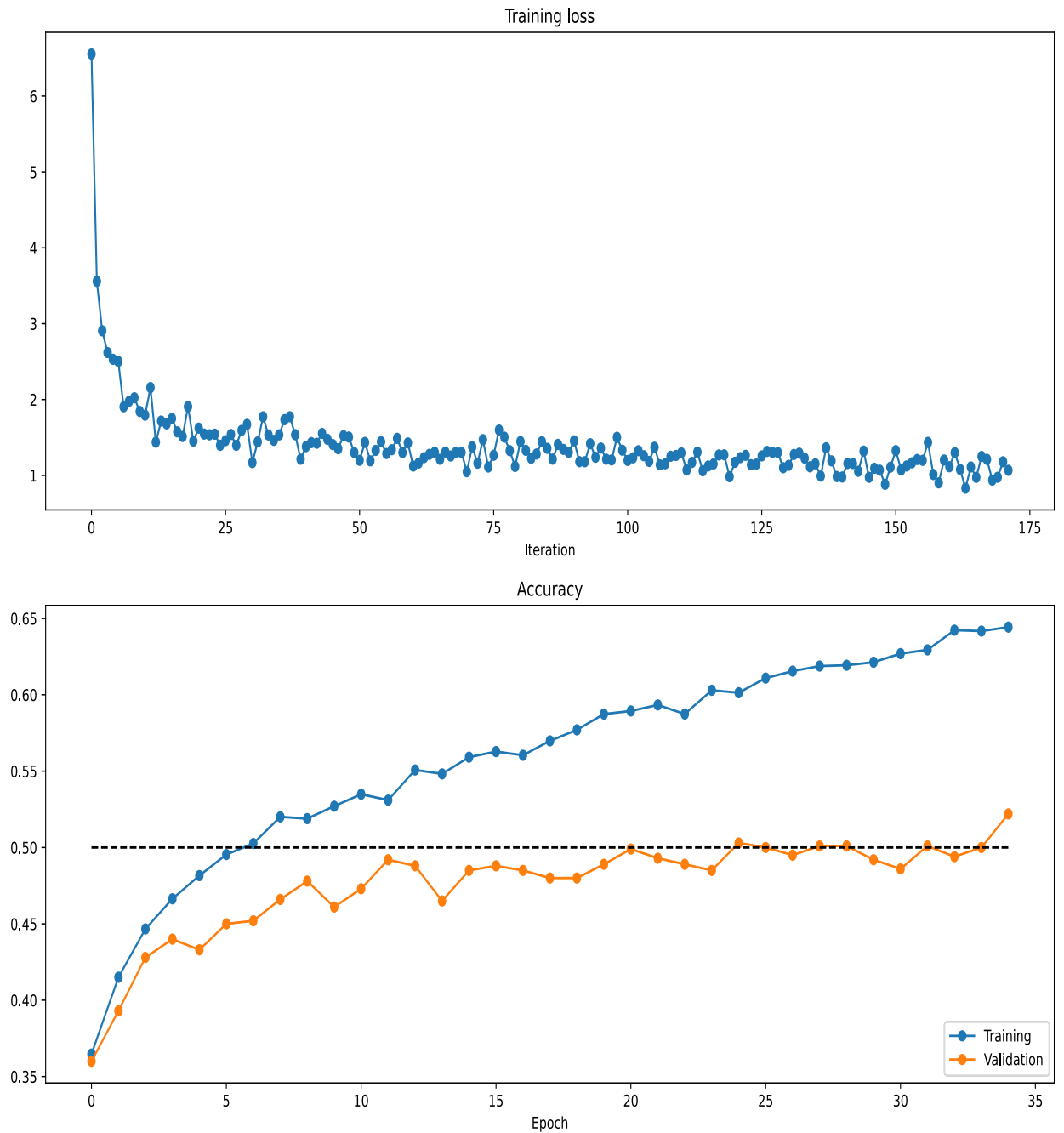
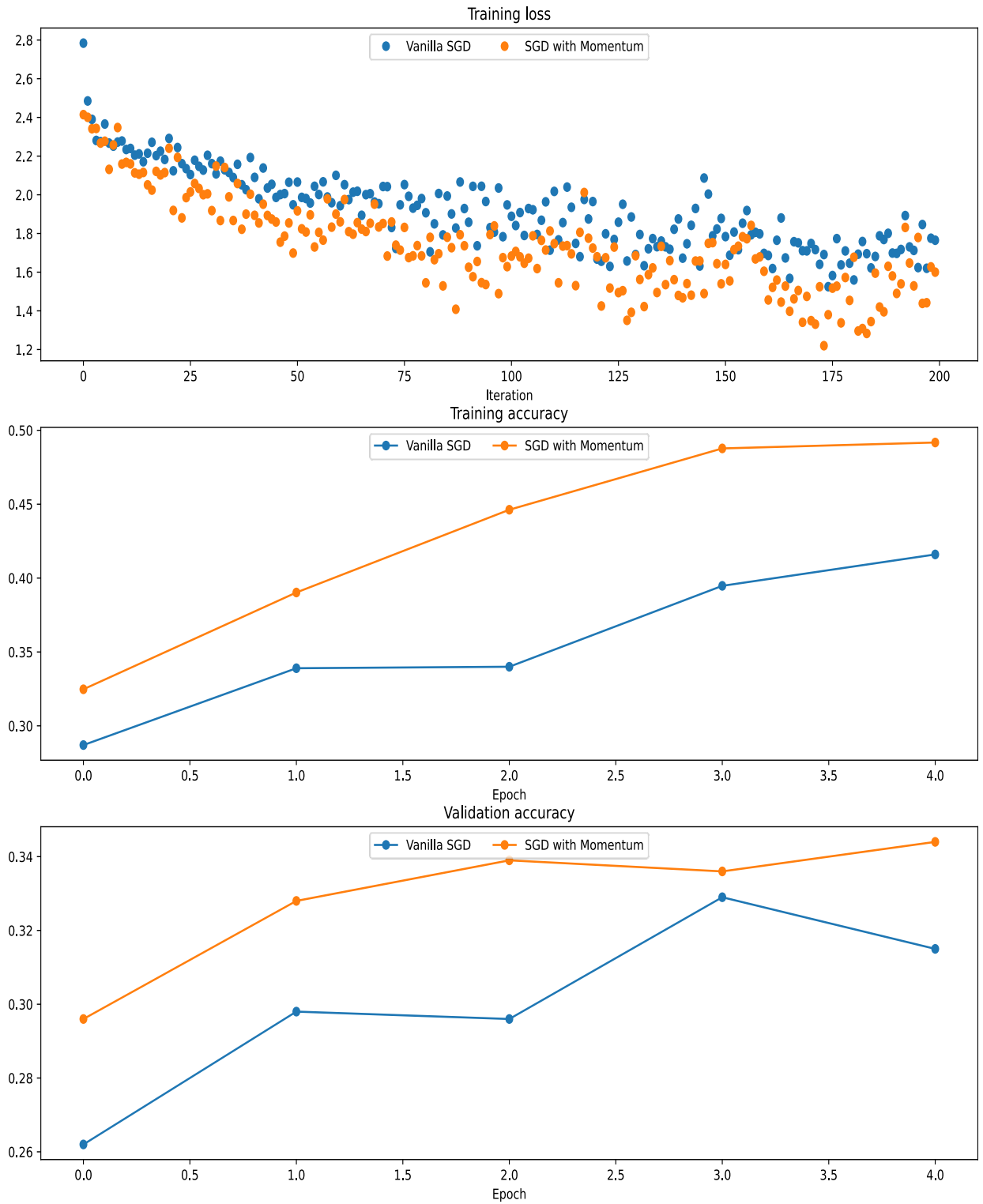


PROBLEM 1

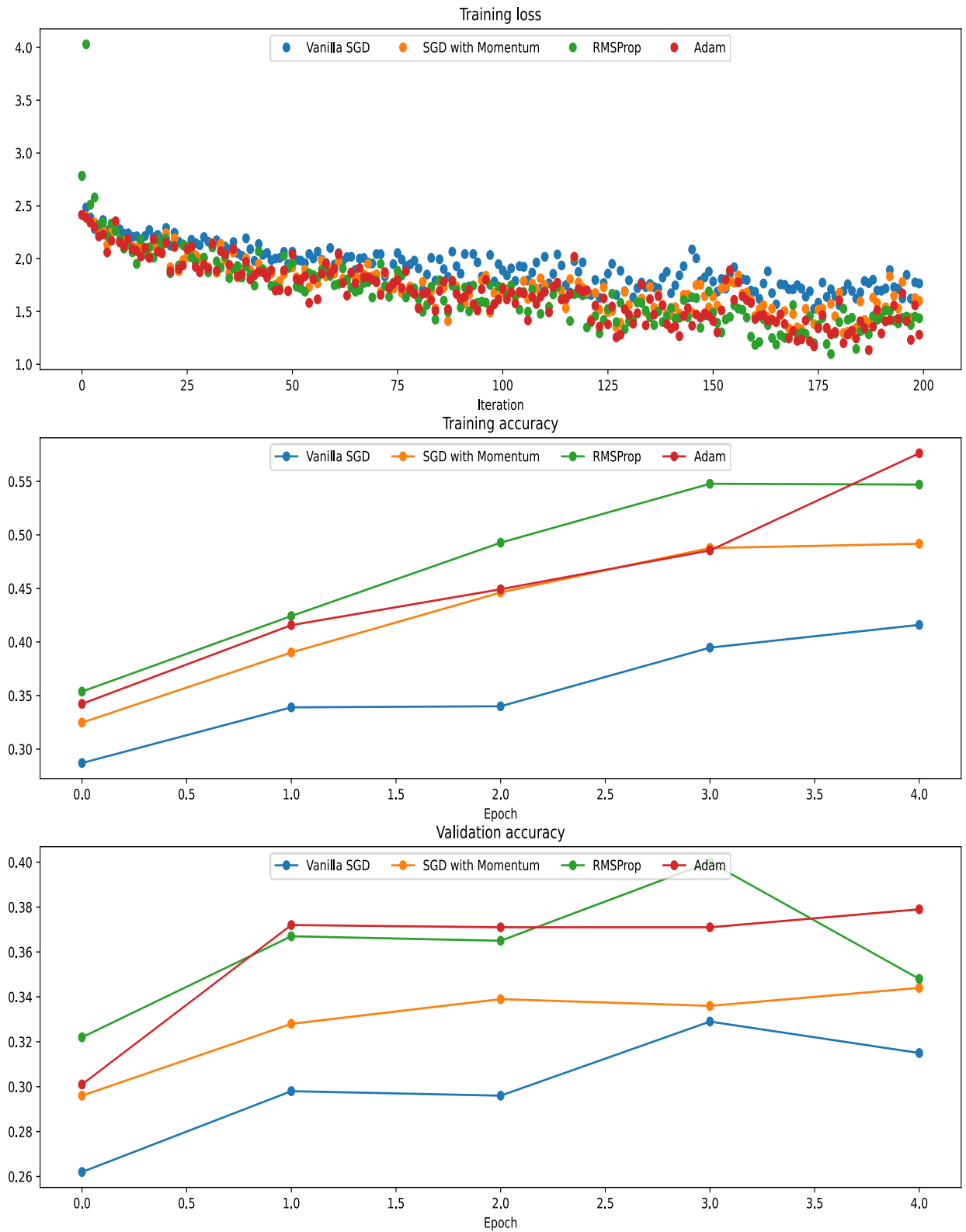
1. Training loss / accuracy curves for the simple neural network training with >50% validation accuracy



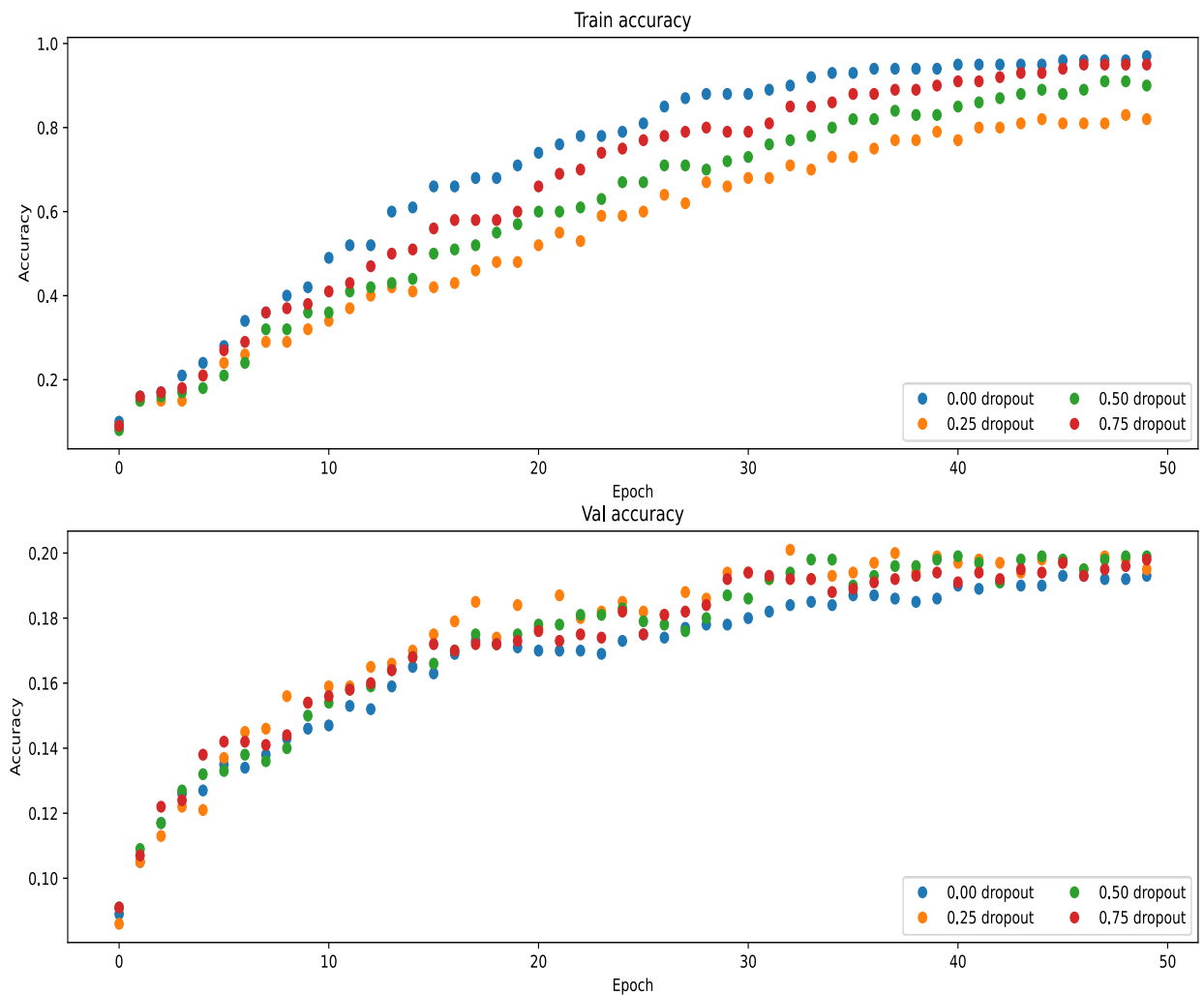
2. Plots for comparing vanilla SGD to SGD + Momentum



3. Comparing different Optimizers" plots



4. Dropout comparison plots



5. Dropout inline question answer

Observation: When the dropout is on the either extremes (close to 0, or 1), then model does well on training data but not so much on validation data in comparison to the models having a reasonable dropout rate, like 0.25 or 0.5. However, those models with reasonable dropout rates gets a lower training accuracy then the other two models.

Explanation: When dropout is 0, the model overfits because there isn't randomization caused by the dropouts. And then, when the dropout rate gets to reasonable range of either 0.25 or 0.5, we get ideal randomization for the model to learn essential features without overfitting. However, when the dropout further goes to other extreme, we get a bit more randomization then ideal and hence the validation accuracy starts dropping.

6. Activation function plot

