

Gradient

I implemented the numerical gradient check to compare my relative error. After receiving the following results:

Relative error W: 3.432027e-03

Relative error B: 4.445333e-07

I concluded that, with my accuracy rising at a comparable rate with what was reported in the assignment, my gradient was at least very close to correct.

Comments

Regularization: It can be a powerful tool to help prevent overfitting while also providing a hyperparameter to optimize in order to reduce the loss (and thus increasing accuracy) when using cross-validation.

Learning rate: It's another hyperparameter that both impacts loss, accuracy and run time. Learning rates are interesting, and the appropriate one varies heavily with the gradient surface the net tries to find a minimum loss within. There are many important concepts of how to implement the learning rate, most commonly and widely used are the adagrad and the adam functions, but of course with a dynamic scale factor that reduces the learning rate once the gradient converges to approach the minimum even more.

Plots

The rest of the report will be filled with the plots I generated during the run. It will have the following format:

1. Output of the whole program
2. General plots about the data
3. Plots about the initial example, using data_batch_1 and data_batch_2
4. Plots about the parameter tests specified in the assignment

Initial Accuracy: 0.12 (note that this was before training, as expected ~10%)

Gradient relative error check:

Relative error W: 3.432027e-03

Relative error B: 4.445333e-07

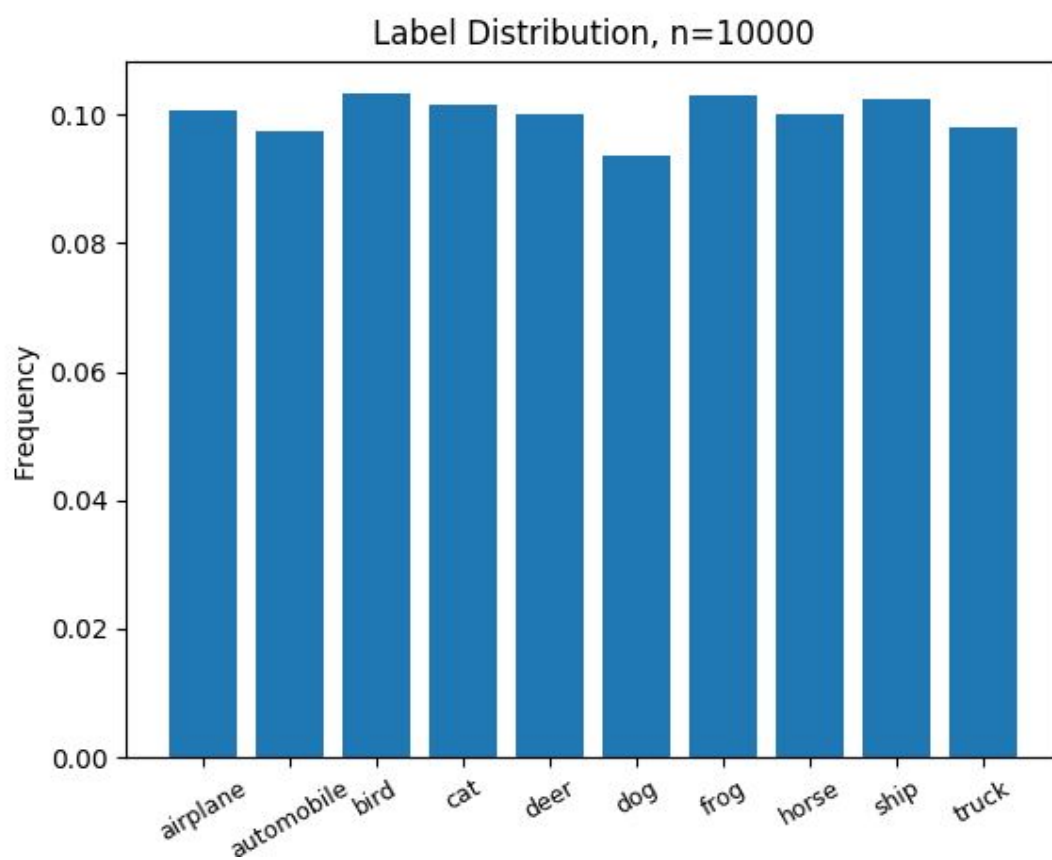
ParamTest1 Test Accuracy: 0.25

ParamTest2 Test Accuracy: 0.39

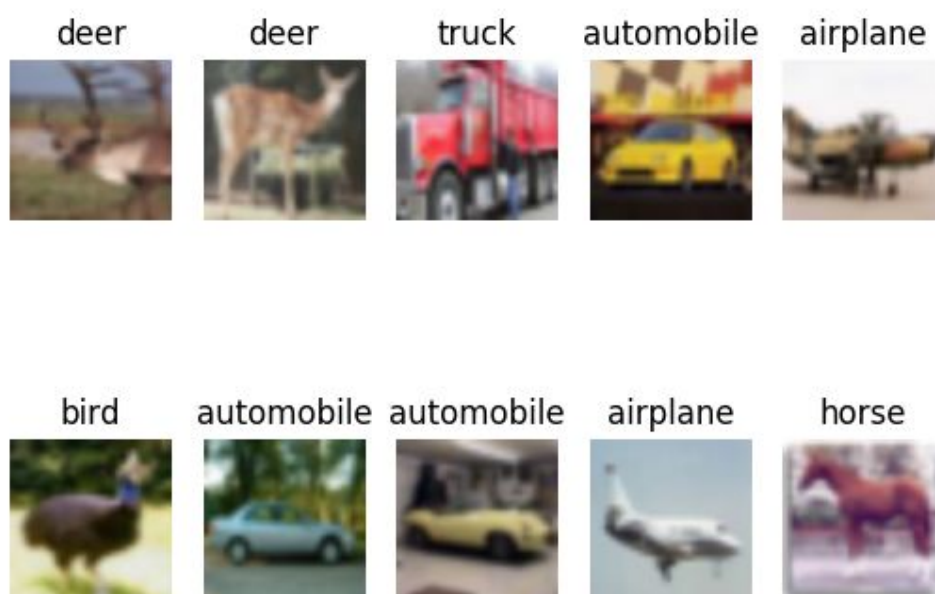
ParamTest3 Test Accuracy: 0.34

ParamTest4 Test Accuracy: 0.24

Note: Original images and output are provided in the folder submitted.



Sample images from Training set



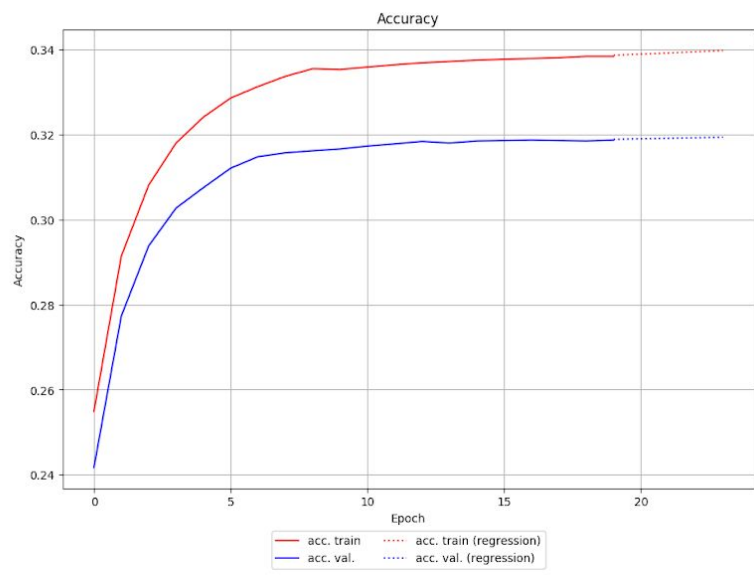
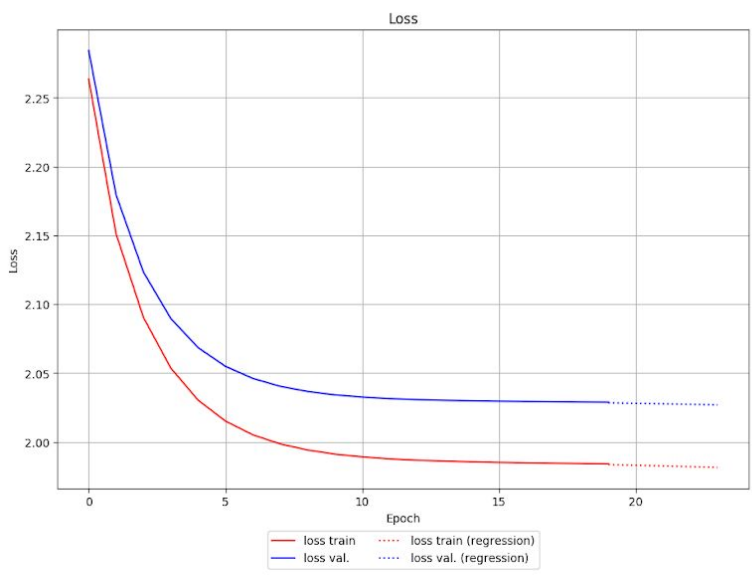
Sample images from Validation set



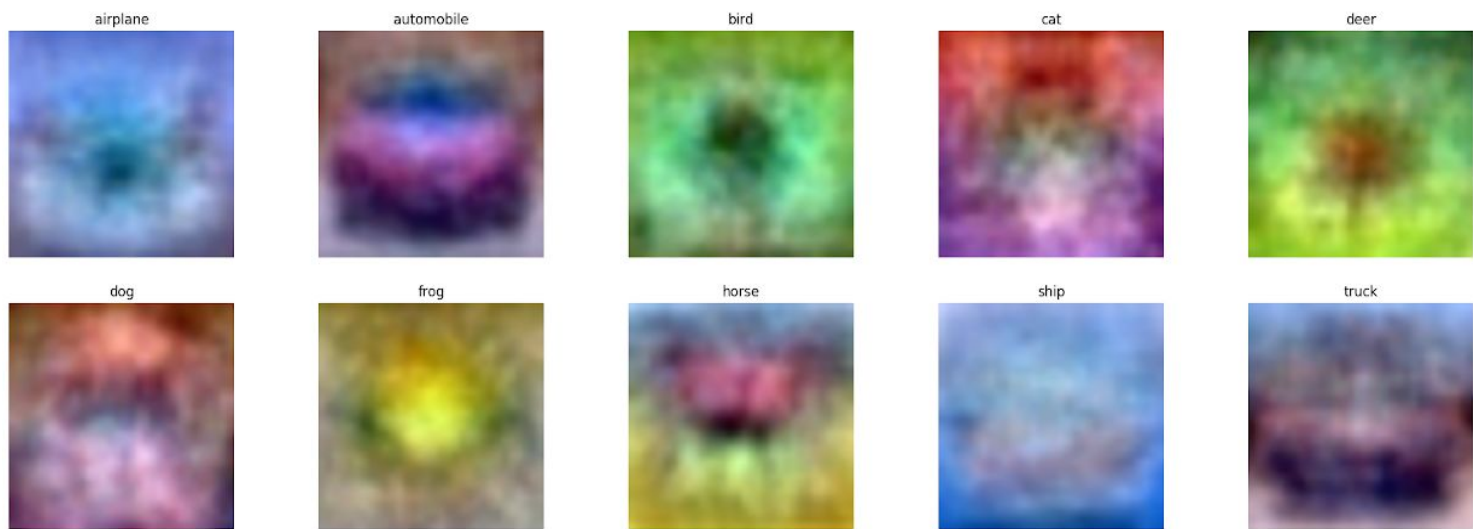
Sample images from Test set



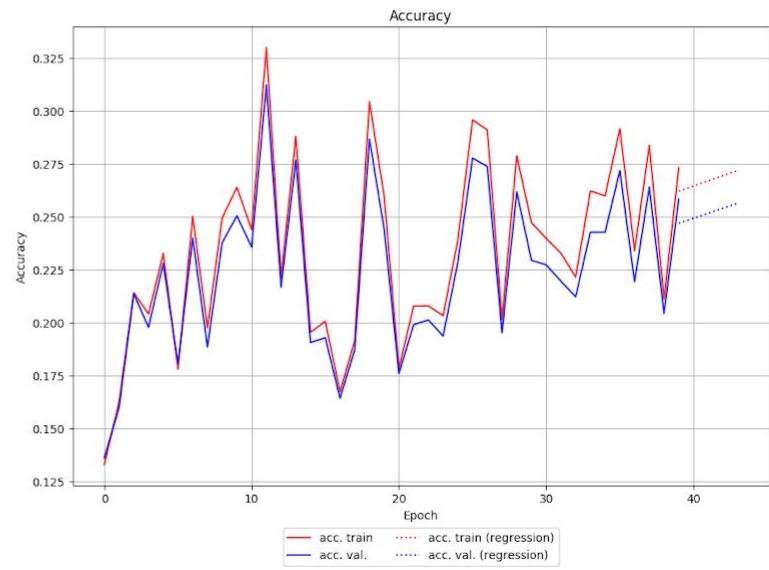
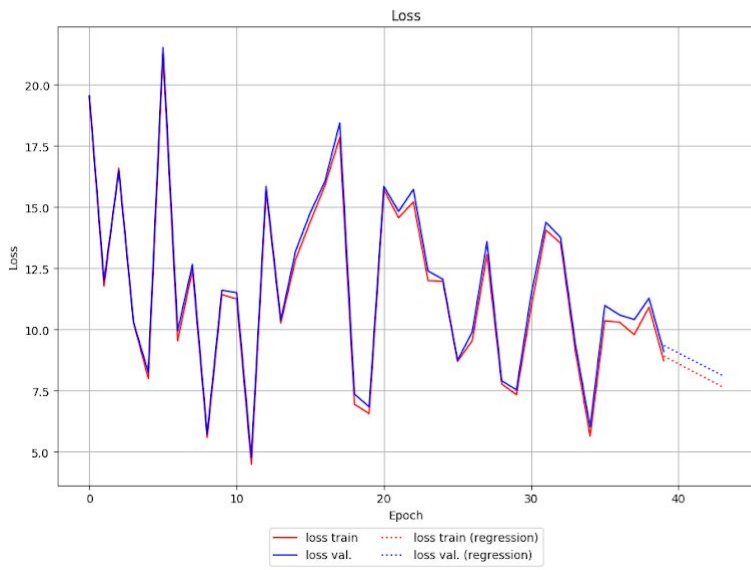
Initial



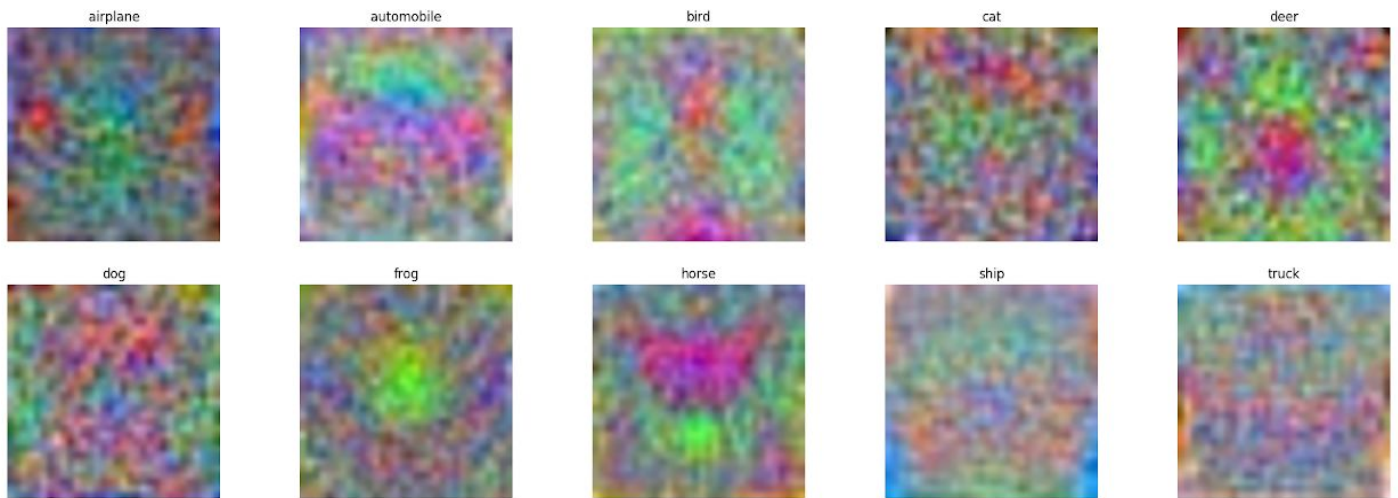
Initial



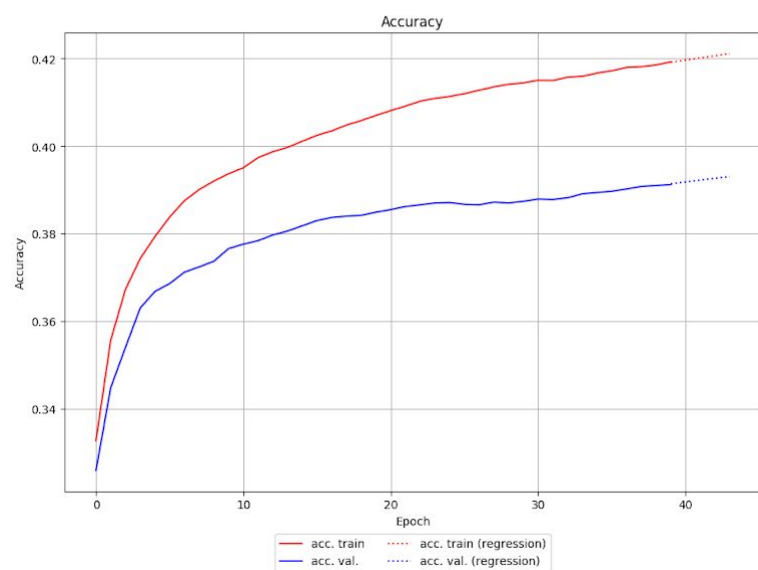
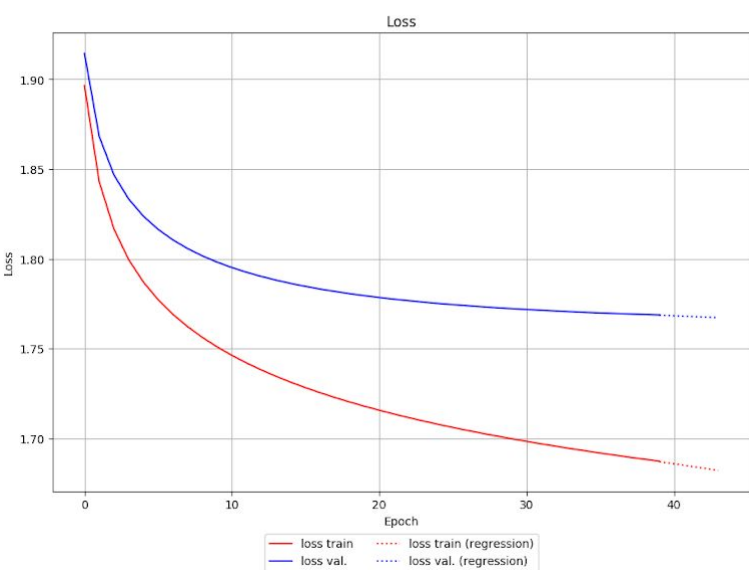
ParamTest1



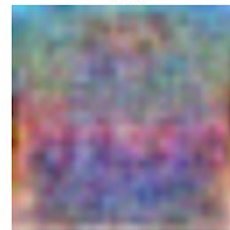
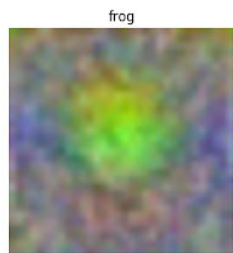
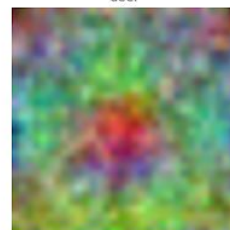
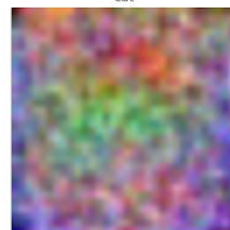
ParamTest1



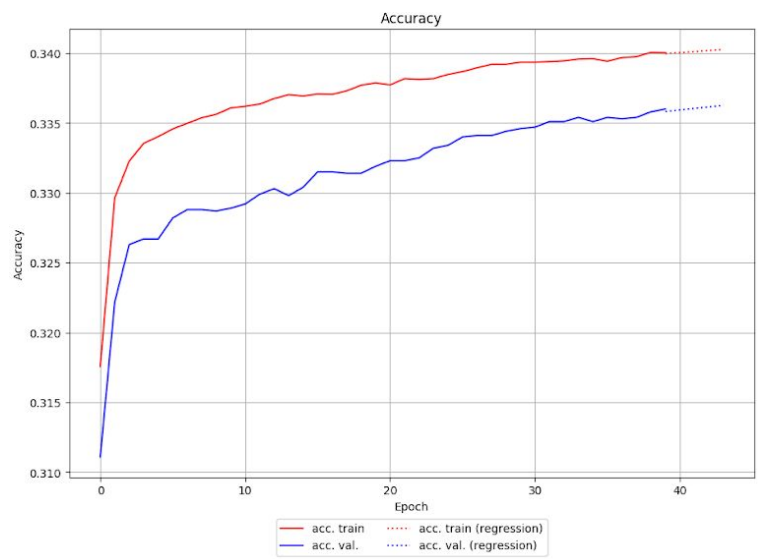
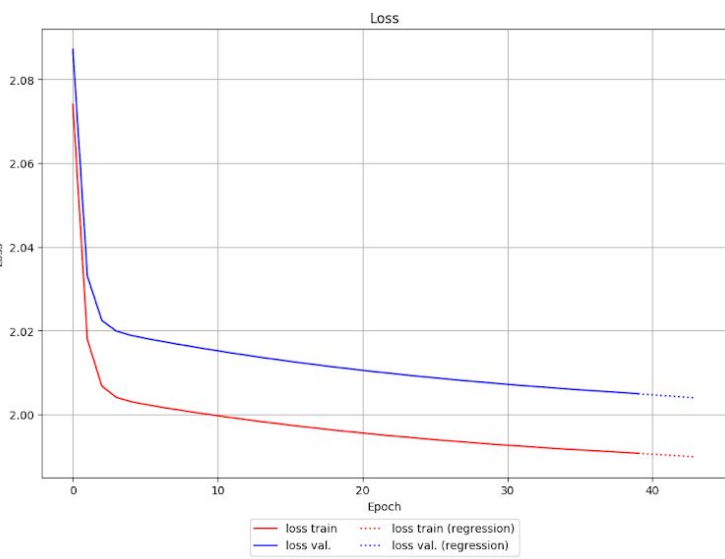
ParamTest2



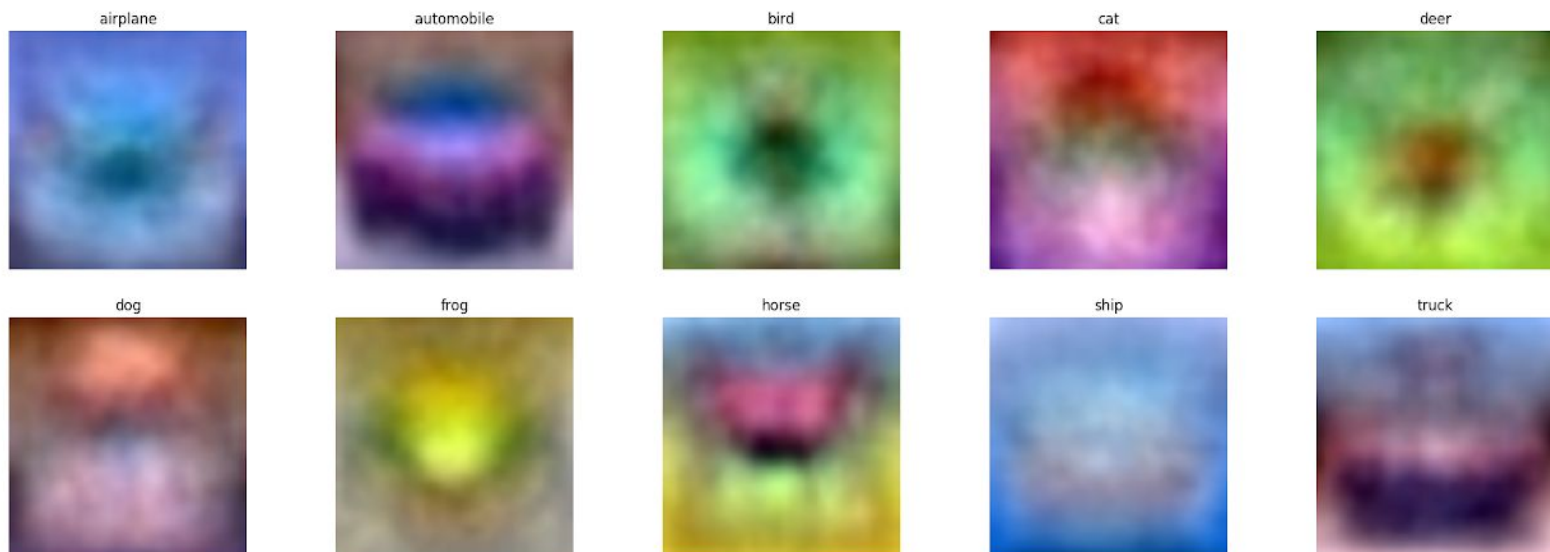
ParamTest2



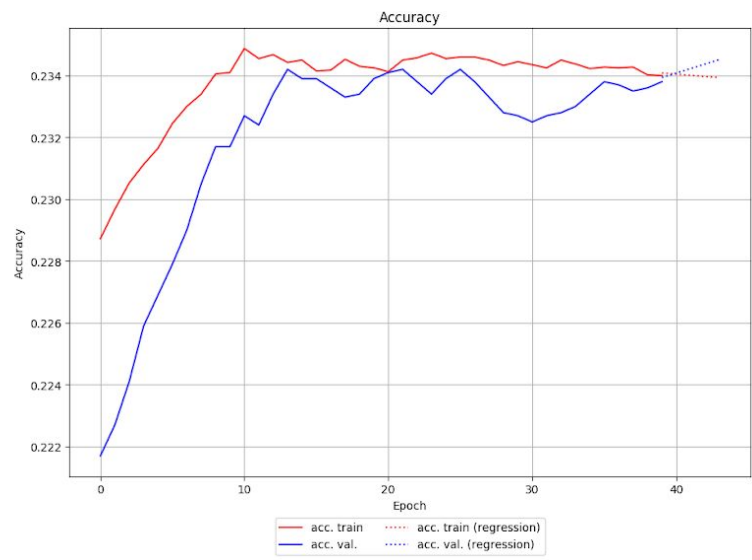
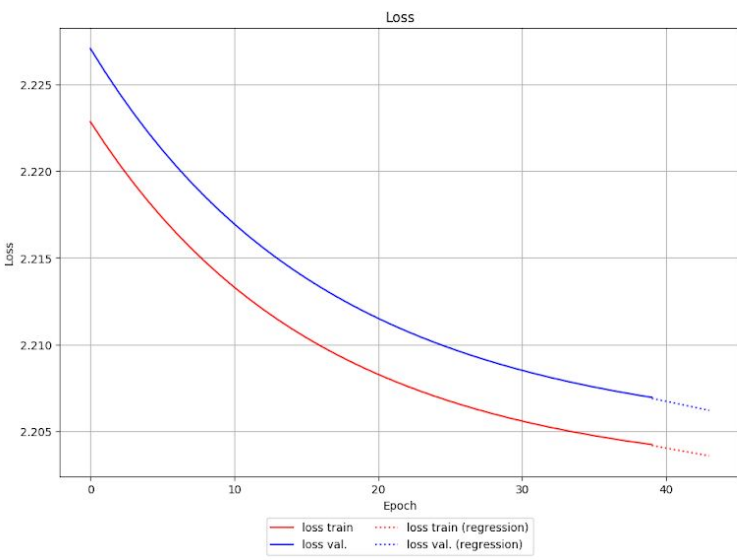
ParamTest3



ParamTest3



ParamTest4



ParamTest4

