Dylan McNamara

ECE5450: Machine Learning

Project Part 2 Write Up

Graphical user interface

Description automatically generatedGraphical user interface

Description automatically generated

1. The first set of figures is the best learning rate found for Adam optimizers (learning rate = .0001) on the left and the best learning rate on SGD optimizer (learning rate = .01) on the right. The accuracy was higher and the loss was lower on the Adam optimizer, so I believe it is the better model to move forward with. I did notice large fluctuations in the accuracy with different learning rates and it seems some model work better with different learning rates.

Graphical user interface, application

Description automatically generatedGraphical user interface, application

Description automatically generated

2. The first set of figures on the left shows the relu selected activation and the sigmoid selected activation on the right. Both of these tests used the Adam optimizer with a learning rate set to 0.0001. The relu activation seemed to remain more consistent on the accuracy, whereas sigmoid seems to oscillate, so I decided to proceed with the relu activation on the model.

Graphical user interface

Description automatically generatedGraphical user interface

Description automatically generated

3. The figure on the left shows with the early stopping strategy implemented, and on the right is without the early stopping strategy. The model used the Adam optimizer, with relu activation, and a set learning rate to 0.0001. The early stopping strategy here saved a round of evaluation, an in the future it may be able to detect when the model has maxed its potential. This feature will be used at will in the future modeling.

Graphical user interface

Description automatically generatedGraphical user interface

Description automatically generated

4. The figure on the left is using data augmentation on the pictures to try to increase the validation accuracy, whereas the figure on the right is using no data augmentation to preprocess the data sets. Both models used an Adam optimizer, relu activation, early stop strategy enabled, and learning rate set to 0.0001. The data augmented accuracy seems to show a greater degree of accuracy overall, so the next models will include data augmentation.

Graphical user interface

Description automatically generatedGraphical user interface

Description automatically generated

5. The image on the left is using a deep set of layers in the model whereas the right is using a wide set within each layer. Using a depth method seems to be more stable than a wide approach but the very wide or very deep approach both led to worse models. I believe a moderate amount of layers with moderate width would provide a better model, so for the CIFAR10 data set, the model will use a slightly wider and deeper model than before. Both models used the Adam optimizer, relu activation, early stop strategy enabled, data augmentation enabled and learning rate set to 0.0001.

CIFAR-10 Resuslts

Graphical user interface, application

Description automatically generated

The final model used Adam optimizer, relu activation, early stop disabled, data augmentation enabled, learning rate set to 0.0001, and an edited set of convolution layers. The model was allowed to run all 100 epochs to see the long term behavior of the model. The validation accuracy seemed to oscillate between 76% and 77% accurate at a maximum. The major issue I believe there might be is that the CIFAR-10 data set does use colored images and I have the same image data generate from the MNIST data set. The layers could most likely be further optimized as well for better performance, as could the image data generator for image augmentation.

In my project I used tensorflow primarily due to my familiarity with it over pytorch. The keras libraries were also very helpful to the implementation of the code. The tensorflow and keras websites provided ample examples that I used as a basis for the neural network, and the tensorflow libraries made creating things like early stop strategies easy. If I was to do this again, I’d most likely spend more time optimizing the data augmentation for colored image in the CIFAR-10 data set.