

CS 475 Machine Learning: Homework 3

Deep Learning

Cassie Parent, Alexandra Szewc
(cparent5, aszewc1)

Instructions

We have provided this L^AT_EX document for turning in homework 3. We give you one or more boxes to answer each question. The question to answer for each box will be noted in the title of the box.

Other than your name, do not type anything outside the boxes. Leave the rest of the document unchanged.

For written answers, replace the `\TextRequired (Place Answer Here)` command with your answer. For plots, replace the `blank.png`

Programming Questions

Question 1.5.8a

Approximate accuracy for random classifier, rounded to whole integer =

17%

Question 1.5.8b

Approximate accuracy for majority-vote classifier, rounded to whole integer =

17%

Question 1.5.9a

Number of Weights =

79000

Question 1.5.9b

Number of Biases =

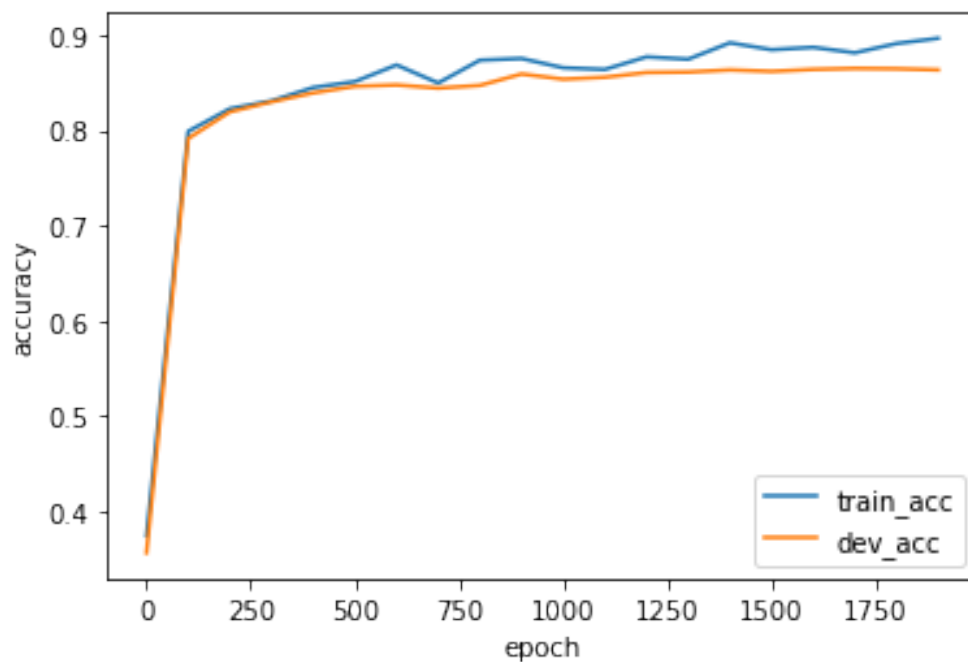
106

Question 1.5.10

Learning rate with best validation accuracy =

0.001

Question 1.5.11 DNN training and validation accuracy plot



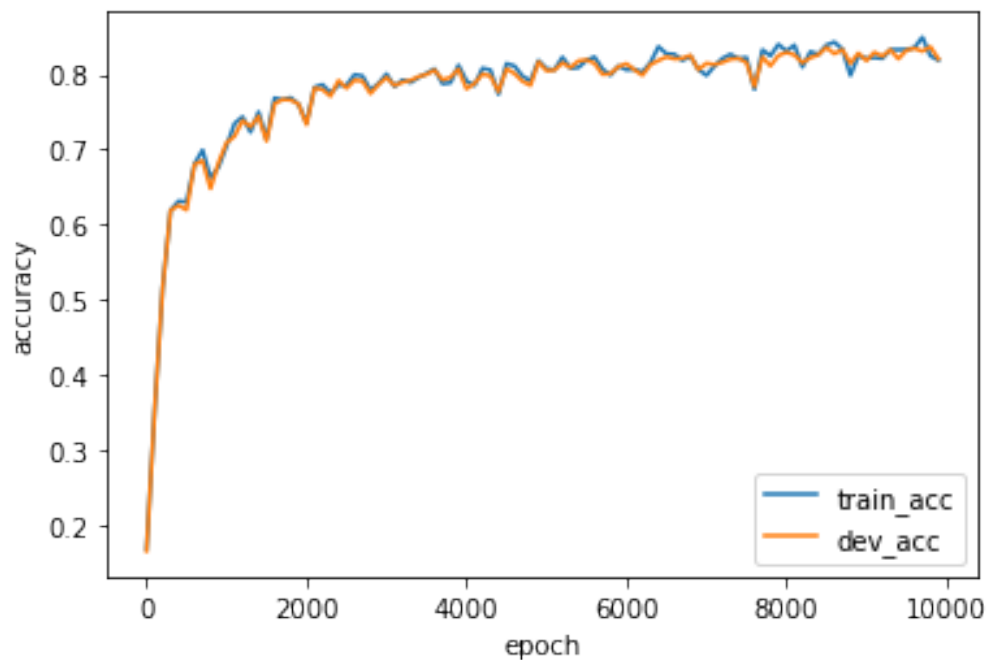
Question 1.5.12. Improve performance with limited dataset?

Apart from using more data, some ways to improve the performance with a limited dataset are to use regularize on network parameters, stop training early when model performance stops improving, or modify the network structure by using dropout.

Question 1.6.2. Convolutional net vs fully connected net?

Fully connected networks are densely connected, while convolutional networks are not as densely connected so not all input points have an effect on the output. This also means that the number of weights per layer much smaller for convolutional networks, which makes them ideal for image data, which has high-dimensional or large input.

Question 1.6.5 CNN training and validation accuracy plot



Question 1.6.6. Describe best parameters

The best training parameters involved a batch size of 100, a learning rate of 0.01, and 10000 epochs. This involved 100 optimization steps and took about 3 minutes to run.

Question 1.6.7

Best validation accuracy, rounded to whole integer =

85%

Question 1.8.4a: Are there any qualitative differences between these sets of images

☒ Yes

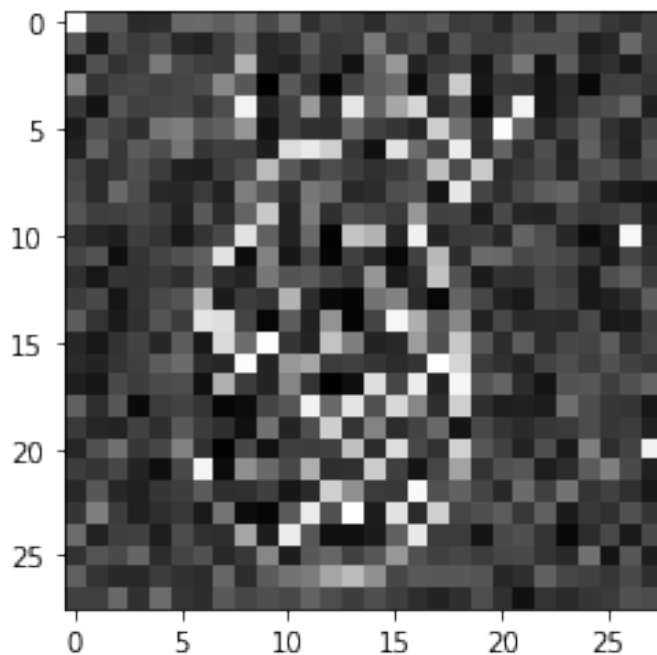
☐ No

Question 1.8.4b: Are the misclassified examples more difficult for you to classify

☒ Yes

☐ No

Question 1.8.6



Question 1.8.7: Does the classifier still classify all 10 images correctly

☐ Yes

☒ No

Question 1.8.10: Does the classifier still classify all 10 images correctly

☐ Yes

☒ No

Question 1.8.11 Flipped images

Failure to vertically/horizontally flip is not necessarily a failure. In this case, robustness to vertical flipping of an apple might lead to confusion between other classes (if they existed) such as a candied apple on a stick. The same applies to horizontal flipping, such as if we had a model that needed to differentiate between asymmetrical classes, such as hands pointing to the left or right.

If we wanted to remain invariant to horizontal flipping, we could augment some of the training data to include instances where inputs of the same class are flipped so that the model can be trained on these types of variations. For example, pointed hands or apples could be flipped horizontally and vertically but still given the same label.