Timothy Horrell

ML Applications in ECE

Homework 5

**Question 1:**

Formula for trainable parameters = input channel x output channel x filter size + bias

The number of parameters in each layer are as follows:

Input layer - None

First Layer - 3 \* 96 \* (11\*11) + 96 = 34,944 parameters

Second Layer – 96 \* 256 \* (5\*5) + 256 = 614,656 parameters

Third Layer – 256 \* 384 \* (3\*3) + 384 = 885,129 parameters

Fourth Layer – 384 \* 384 \* (3\*3) + 384 = 1,327,488 parameters

Fifth Layer – 384 \* 256 \* (3\*3) + 256 = 884,992 parameters

Total parameters: 3,747,200 parameters

AlexNet\_CNN Summary:

A screenshot of a computer program

Description automatically generated

The parameters line up exactly as expected from my calculation in part A.

**Question 2:**

The CNN was trained with mini-batch SGD, with a batch size of 5. The validation loss was evaluated after each epoch, while training loss was evaluated after each batch.

The \_\_ network gives the best prediction accuracy on the testing dataset. The activation functions have their own strengths and weaknesses. Because the sigmoid is bound to 0 and 1, it is resistant to ‘blowing up’ the activation. The Relu however holds on to attributes of parameters far from the y-axis. For example, far reaching values on the sigmoid x-axis would equal 0.001 or 0.00001, which are both essentially 0.

Based on these results, the sigmoid function took longer to learn, and didn’t end up performing as strongly as the ReLU function. Additionally, it had longer runtime, meaning the ReLU is stronger in all facets here.

ReLU Testing Accuracy:

A black screen with white text

Description automatically generated

ReLU Training Loss:

A graph of a training

Description automatically generated

ReLU Validation Loss:

A graph with numbers and a line

Description automatically generated

Sigmoid Testing Accuracy:

\*\*The sigmoid accuracy goes from epoch 0 to 4, like ReLU. This was a typo in my code.

A black screen with white text

Description automatically generated

Sigmoid Training Loss:

A graph of a person

Description automatically generated

Sigmoid Validation Loss:

A graph of a person

Description automatically generated