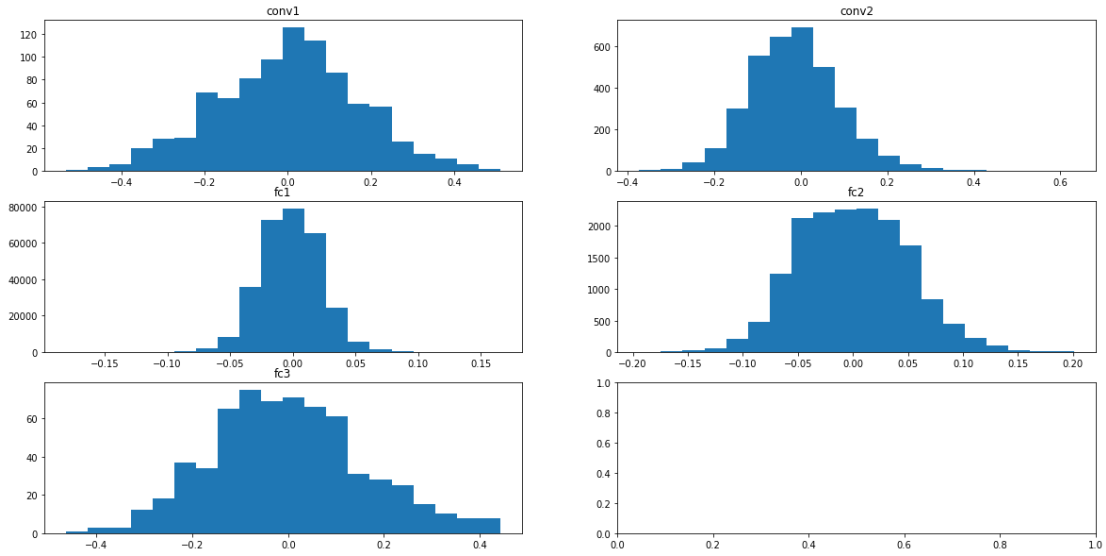
Initial Setup

Accuracy: 61.2%

Question 1.1



Question 1.2

Layer	Actual range	3-sigma range
Conv1	1.0433	1.0238
Conv2	1.0051	0.6252
Fc1	0.3464	0.1418
Fc2	0.3948	0.2995
Fc3	0.9068	0.9607

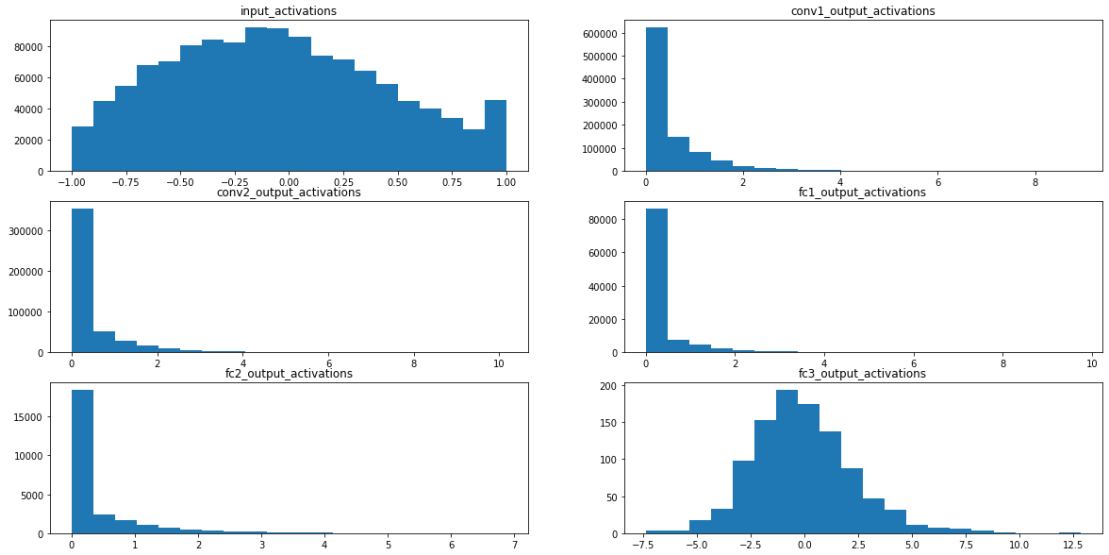
Fc3 is the 3-sigma range larger than the actual range.

Other layers are the 3-sigma ranger smaller than the actual range.

Question 2.2

Accuracy: 60.98%

Question 3.1



Question 3.2

Activation	Actual range	3-sigma range
input_activations	2.0	2.98
conv1_output_activations	8.93	4.18
conv2_output_activations	10.17	4.19
fc1_output_activations	9.74	3.47
fc2_output_activations	6.88	4.20
fc3_output_activations	20.23	14.22

Question 4.1

(a) n_Wconv1*W_conv1*n_In*In*n_Outconv1=n_Wconv1*n_In*n_Outconv1*Out_conv1

(b) n_Wconv2*W_conv2*
n_Wconv1*n_In*n_Outconv1*Out_conv1*n_Outconv2=n_Wconv2*
n_Wconv1*n_In*n_Outconv1*n_Outconv2*Out_conv2

Question 4.2

Accuracy: 61.06%

Question 4.4 (Bonus)

1. Use unsigned 8-bit number [0, 255] for quantized activation. Because the activation after ReLU will become non-negative value, so unsigned number can quantity activation better.
2. Use asymmetric quantization, move the zero point to the mean of the activation. By this method, the signed 8-bit number can cover larger range of the activation.

Question 5.1

Scale should be: nW*nIn*nOut

Accuracy: 59.21%

Question 5.2

Accuracy: 62.37%