Lab 5

Question 1:

gradient = tape.gradient(loss, parameters)

Question 2:

100480 = Neurons in first layer * (the flattened input shape + the line intercept/constant) =

= 128 * (28 * 28 + 1)

Question 3:

Batch size 1 (Column 1):

60000/60000 [==============================] - 135s 2ms/step - loss: 0.5278 - accuracy: 0.8122 313/313 [==============================] - 1s 2ms/step - loss: 0.5053 - accuracy: 0.8182 Test accuracy: 0.8181999921798706

Batch size 10 (Column 2):

6000/6000 [==============================] - 14s 2ms/step - loss: 0.4800 - accuracy: 0.8245 313/313 [==============================] - 1s 2ms/step - loss: 0.4166 - accuracy: 0.8450 Test accuracy: 0.8450000286102295

Batch size 100 (Column 3):

600/600 [==============================] - 3s 4ms/step - loss: 0.5263 - accuracy: 0.8159 313/313 [==============================] - 1s 2ms/step - loss: 0.4538 - accuracy: 0.8384 Test accuracy: 0.8384000062942505

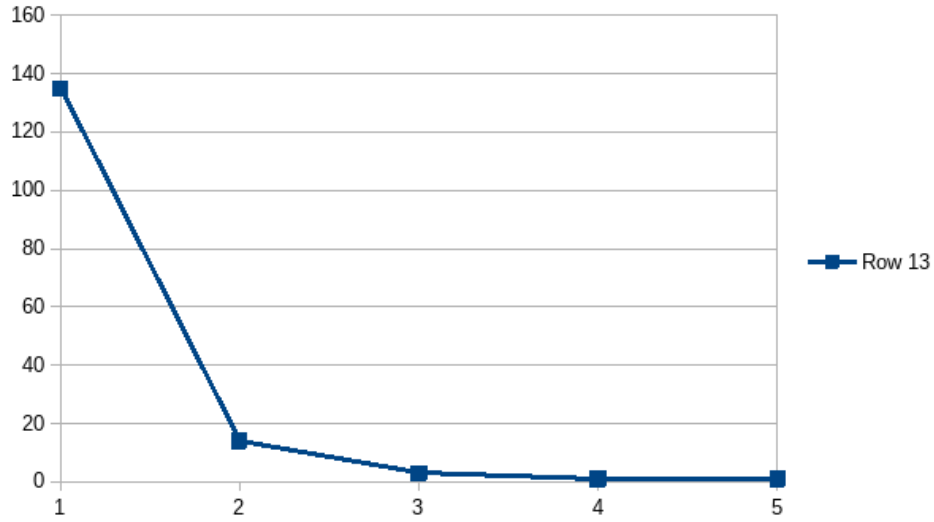
Batch size 1000 (Column 4):

60/60 [==============================] - 1s 16ms/step - loss: 0.8351 - accuracy: 0.7194 313/313 [==============================] - 1s 4ms/step - loss: 0.5346 - accuracy: 0.8176 Test accuracy: 0.8176000118255615

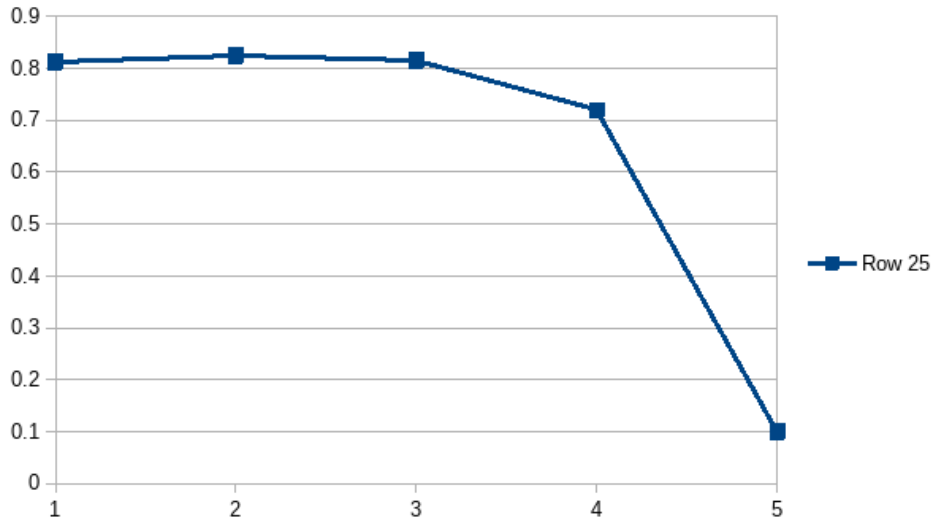
Batch size 60000 (Column 5):

1/1 [==============================] - 1s 1s/step - loss: 2.4001 - accuracy: 0.1007 313/313 [==============================] - 1s 2ms/step - loss: 2.1806 - accuracy: 0.3012 Test accuracy: 0.3012000024318695

Training time:



Accuracy:



Batch size 1000, gives the best accuracy for the time spent.