Model loss for Architecture: adam optimizer, 2 layers, 256.0 nodes, 200 epochs, 0.001 learning rate, 128 batch size