Model loss for Architecture: adam optimizer, 2 layers, 128.0 nodes, 50 epochs, 0.001 learning rate, 128 batch size