

Loss function and accuracy, AdaGrad SGD w/500 epochs - no gradient optimizer

