Training MSE for SGD with 500 epochs

