Training MSE for SGD with 5000 epochs

