Stochastic Gradient Descent Learning Rate = 0.5 Hidden Units = 100

