

AdaGrad MSE stochastic gradient descent $\gamma=0.4$ 