

AdaGrad MSE stochastic gradient descent with momentum

