

正确 Try using Adam 10° Which of the following statements about Adam is False? Adam combines the advantages of RMSProp and momentum 0/1分 We usually use "default" values for the hyperparameters eta_1,eta_2 and arepsilon in Adam ($\beta_1 = 0.9, \beta_2 = 0.999, \varepsilon = 10^{-8}$) Adam should be used with batch gradient computations, not with minibatches. The learning rate hyperparameter lpha in Adam usually needs to be tuned. 这个选项的答案不正确