神经网络第七章 2022年7月17日 星期日 luning Process (调试如理) D. B. B. B. S. S ... (2か多) (adam) # layers # hidden units mini-batch size the learning rate decay (8) 取点, 看 36 号 (A) 需要确定的哪个超渗的面数果长! (进行点抽样判断) Using an Appropriate Scale to pick Hyperparameters 取一它尺度来限制超参抽样范围 # layers 1: 2-4 7,3,4 חססי0 0.00 0.00 0.0 由 python 表 寸: Y=-4* NP. random. rand() K=10° re[-4,0], 2 e[104,1] 0 椒重设置 B-0.9 .. - 0.999 10 (10倍敏速) 1000 (1000倍敏感) 实操: pandas VS. By model I. 只能进约一个Midel Caviar L鱼子酱) Strategy panda strategy 少让で自己配 (pandas) > 时时调参 CPU/GPU TIMS H (社吃佛留-分 (baby sitting) Normalizing Activations in a Network. M= M ZXW) χ x = x - M62 = 1 27 X(1)2 X3 X= X/62 廊世调 a^{cz}使 U^{cz}, b^{cz}能更 快调参 XI X 2 Ά3 in plementing Botten Norm 7 Z (m) M= # Z; Z() (Zi-M) Z (Zi-M) Z("), ..., Z(m) Z(i) = Z(i)-M 16249 一> 芝(1) = 孔云(1) +B (可以膻型中更好了, B) 当 8= 1632 B=11. 附, 长, = Zi Fitting Bates Norm into en neural network Adding Batch Norm to a network > 50 ath_ X2. X WIN, bin Z EIN ATIN Z IN DATINE g ((Z iii) W 165 Z IN BIN, ATIN D at BII) = BTI) - dolBTI) Batch Norm at test time M= # Z Zin 62 = # Z (Z(i)-N)2 Z (1) = Z(1) - M 16249 Z(i) = JZ(i) norm +B $t = e^{(2t)}$ $a^{t11} = e^{2t}$ $a^{t11} = e^{2t$ Softmax Regressim 犯到多种回旧结果 Softmax 决集此界. 有多种变型分 ~P(bc(x) 有,a^{t1]}:g^{T1]}(Z^{T1]}) $= \begin{bmatrix} e^{5}/e^{5}+e^{7}+e^{7}+e^{7} \\ e^{7}/e^{5}+e^{7}+e^{7}+e^{7} \end{bmatrix} = \begin{bmatrix} 0.842 \\ 0.042 \\ 0.042 \\ 0.002 \\ 0.002 \\ 0.002 \end{bmatrix}$ Softmax中历根本色数:工(y,y)=-空yjly分 雨亲用梯度下降. dz1 = ŷ - y-深度学习框架 ° Caffe / Gaffe Z · CNTK · Keras · Lasagne · maxnet · Paddle Paddle · Tensor Flow

"Theano

lensor Floru

· Torch