2018-4-1 学习笔记

**1.DNN:**

Fprop and bprop like shallow NN.

**2.hyper parameters(**

learning rate/iterations/hidden layers/hidden units/activation function …)

**3.Train/dev/test/ data**-----70/30 or 60/20/20

**4.Bias/Variance**

**5.Regularization**

**6.Vanishing/exploding gradients**

Initialize , if g(z)=Relu use

**7.Gradient checking for bprop**