Hyperparameter tuning

Using an appropriate scale to pick hyperparameters

Picking hyperparameters at random

$\rightarrow n^{[l]} = 50 , ..., 100$

$[\times \ \times_{\times \times} \ \times_\times \ _\times \ ^\times \ \times]$

50   100
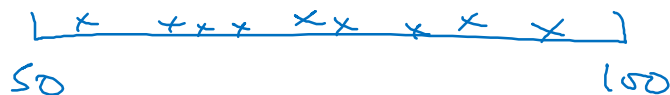
$\rightarrow$ #layers $L: \quad 2 - 4$

$2 \ , \ 3 \ , \ 4$

Appropriate scale for hyperparameters

$\alpha = 0.0001 \ , \ ....., \ 1$

wrong:

0.0001

0.1 10%   90%

right:

0.0001   0.001   0.01   0.1   1

10^{-4} 10^a

$a = \log_{10} 0.0001$
$= -4$

$r = -4 * np.random.rand()$ $r \in [-4, 0]$

$\alpha = 10^r$ $10^{-4} \ ... \ 10^0$

10^0 10^b $b = \log_{10} 1$
$= 0$

$10^a .... 10^b$ $r \in [a, b]$ $\alpha = 10^r$
$[-4, 0]$

Andrew Ng
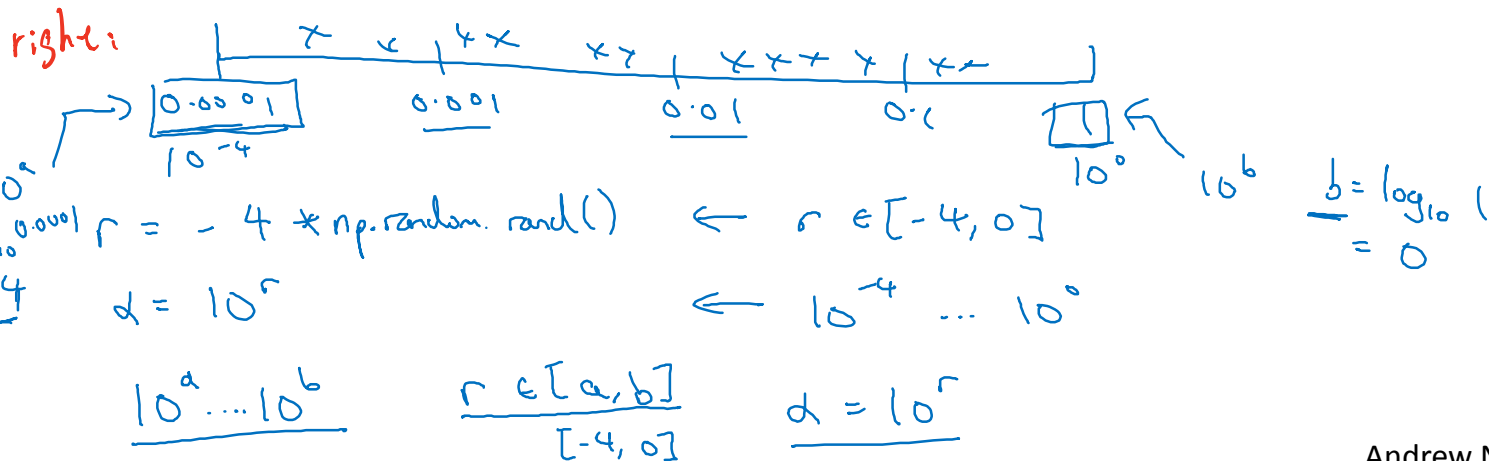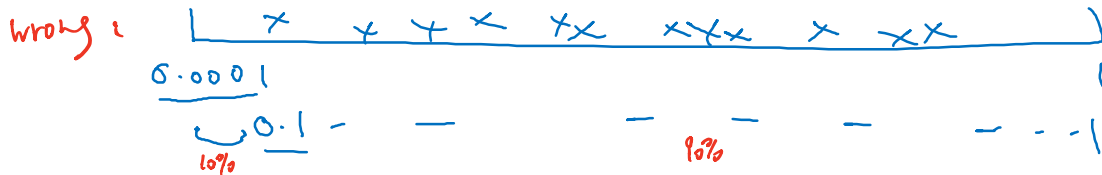
Hyperparameters for exponentially weighted averages

$\beta = 0.9 \quad \cdots \quad 0.999$

$\quad\quad \updownarrow \quad\quad\quad\quad\quad \Downarrow$

$\quad\quad 10 \quad\quad\quad\quad\quad 1000$
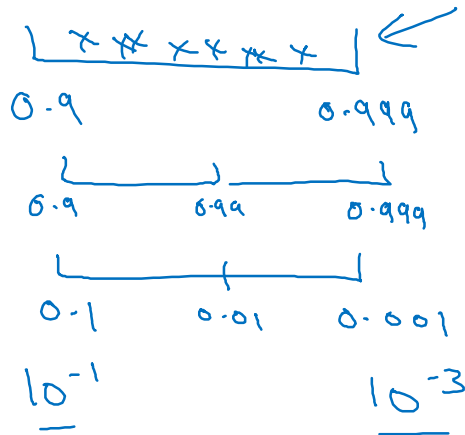
$1-\beta = 0.1 \quad \cdots \quad 0.001$

$\beta: 0.9000 \rightarrow 0.9005 \; \} \sim 10$

$\beta: 0.999 \rightarrow 0.9995$

$\quad\quad \sim 1000 \quad\quad \sim 2000$

$\dfrac{1}{1-\beta}$

$0.9 \quad\quad\quad\quad 0.999$

$0.9 \quad\quad 0.99 \quad\quad 0.999$

$0.1 \quad\quad 0.01 \quad\quad 0.001$

$10^{-1} \quad\quad\quad\quad 10^{-3}$

$r \in [-3, -1]$

$1-\beta = 10^r$

$\beta = 1 - 10^r$

Andrew Ng