



deeplearning.ai

random sampling이 uniformly random을 얘기하는 것은 아님.
appropriate scale을 선택하는 것이 중요하다.

Hyperparameter tuning

Using an appropriate
scale to pick
hyperparameters

Picking hyperparameters at random

이런 경우 uniformly random이 적절하다.

$$\rightarrow n^{\text{test}} = 50, \dots, 100$$

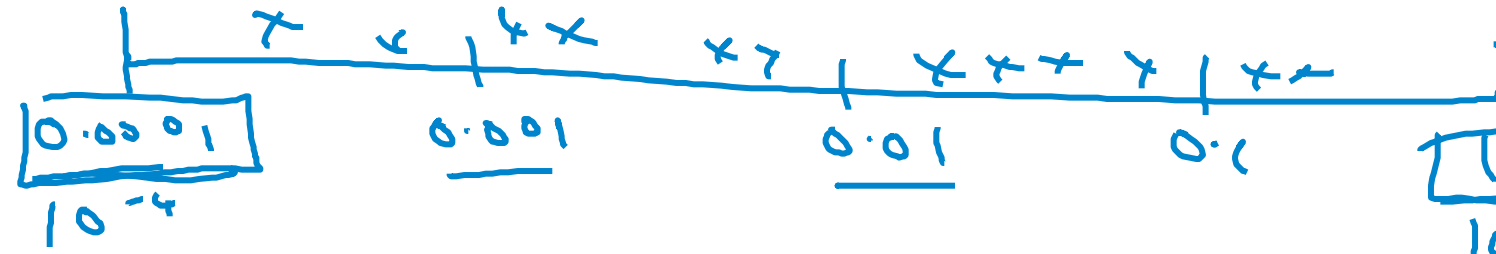
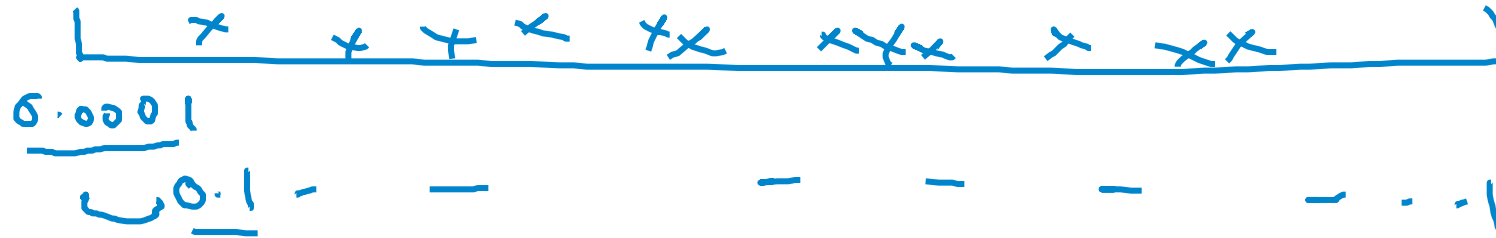


$$\rightarrow \# \text{layers} \quad L : 2 - 4$$

$$2, 3, 4$$

Appropriate scale for hyperparameters

$$\alpha = 0.0001 \dots, 1$$



$a = \log_{10} 0.0001 = -4$
 $r = -4 * \text{np.random.rand}()$
 $\alpha = 10^r$
 $10^a \dots 10^b$
 $r \in [a, b]$
 $r \in [-4, 0]$
 $\alpha = 10^r$
 $b = \log_{10} 1 = 0$
 a, b 사이에서 uniform sampling 해서 그 값을 사용

Hyperparameters for exponentially weighted averages

$$\beta = 0.9 \quad \dots \quad 0.999$$

\downarrow \downarrow
 10 1000

$$1 - \beta = 0.1 \quad \dots \quad 0.001$$

이 경우에는 10일 차이지만

$$\beta: 0.999 \rightarrow 0.9995 \quad \sim 10$$

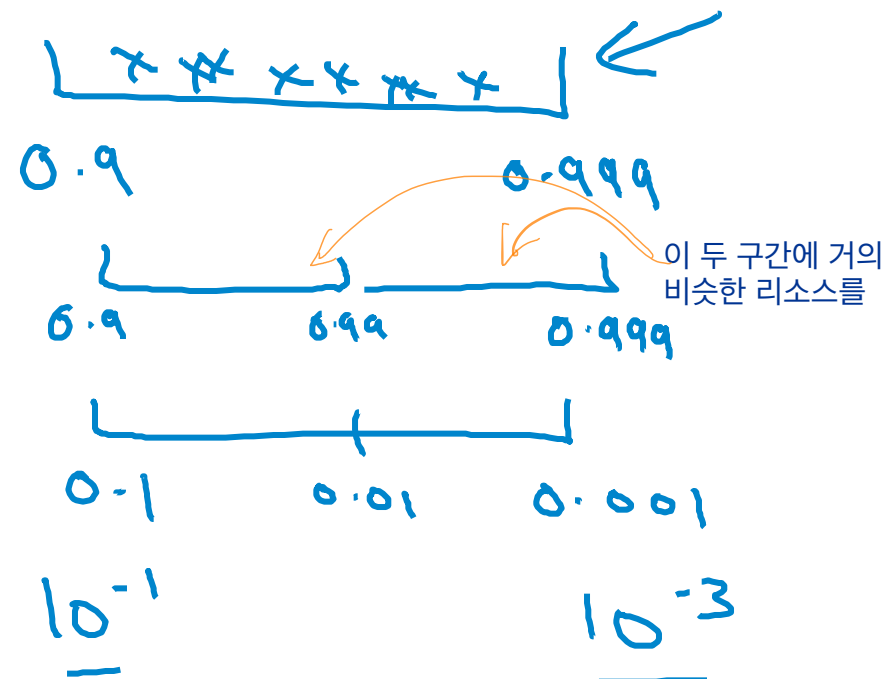
이 경우에는 1000일 차이이다.

$$\beta: 0.999 \rightarrow 0.9995$$

~ 1000 ~ 2000

$$\frac{1}{1 - \beta}$$

우리가 사용하는 수식이 이렇기 때문에 beta가 1에 가까울때 작은 변화에 매우



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

$$1 - \beta = 10^r$$

$$\beta = 1 - 10^r$$