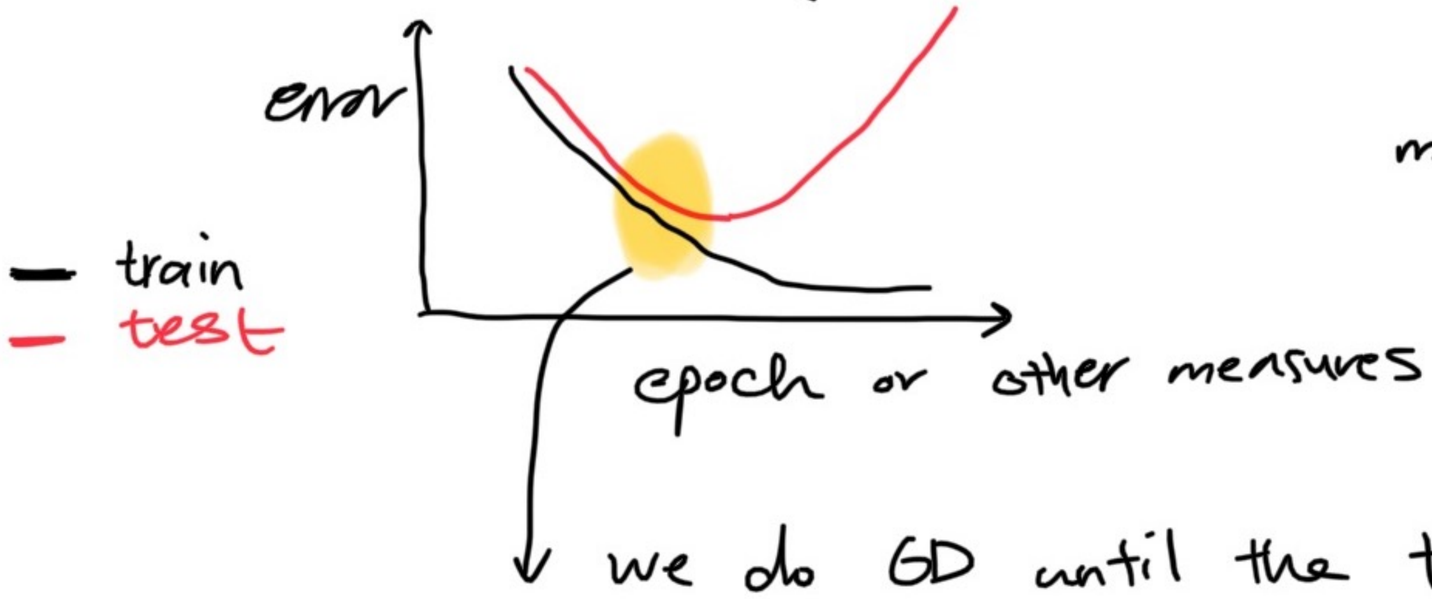
Basic ML KeyWords

underfitting = error due to bias (high bias)

overfitting = error due to variance (high variance)

$$\text{indicator} = \begin{cases} \text{test error : large} \\ \text{train error : tiny} \end{cases} \longrightarrow \text{overfit}$$



model complexity graph

we do GD until the testing error is decreasing. Once it starts to increase, we stop GD. → Early Stopping

Batch Gradient Descent → for all the data

Stochastic Gradient Descent → make small batches of data

Learning Rate Decay:

rule of thumb: if the model is not working increase the rate.