



deeplearning.ai

# Regularizing your neural network

---

## Other regularization methods

# Data augmentation

Say we need  
more Images  
of cats



Synthetic data  
→  
Generation



Flip the Image  $\Rightarrow$  more data

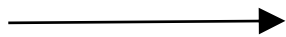
Random transformation of the Image  $\Rightarrow$  more data



- Not as good  
as Independent  
new data, but  
much less  
expensive!

$\Rightarrow$  What we imply is, if something is a cat  $\Rightarrow$  Its "mirror/transformed" Image is also a cat

4



4

4

4

$\Rightarrow$  More data  $\Rightarrow$   $\downarrow$  variance  $\Rightarrow$  Regularization

# Early stopping

Disadv

In ML, we do 2 things

- optimize cost func  $J(w,b)$  (Gradient descent)
- then ↓ overfitting (Regularization-Ridge/Lasso)

These are done separately  
"orthogonalization"

In DNN, we couple the two like this eg →

- Initially almost all "w's" in all layers  $\approx 0$  or some small value
- As we start training, we update w values
- By the end of training, very few "w" values = 0  
⇒ we've used up all w's with some or the other value
- would be good to have a set of w's that's only 60% non zero  
w's being 0 is the same as Regularization ( $\frac{1}{m} \sum w^2$ ) ⇒ If we don't have some of the w values  
 $\frac{1}{m} \sum w^2 \downarrow$

