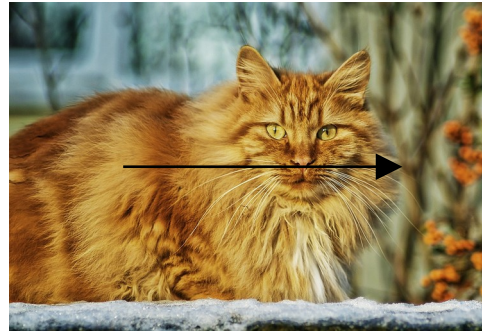deeplearning.ai
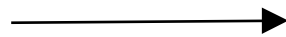
Regularizing your neural network

Other regularization methods

Data augmentation

Early stopping

Orthogonalization.

Optimize cost function J
— Gradient, ...

Not overfit.
— Regularization, ...

l_2

$J(w,b)$

dev set error

trainy error or J

iterations

$w \approx 0$

mid-size $\|w\|_F^2$

large w