



deeplearning.ai

Setting up your goal

When to change dev/test sets and metrics

Cat dataset examples

Metric + Dev : Prefer A
You / users : Prefer B

Metric: classification error

Algorithm A: 3% error → pornographic

✓ Algorithm B: 5% error

Error : ~~$\frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{y_{pred}^{(i)} \neq y^{(i)}\}}$~~

$\sum_{i=1}^n w^{(i)} \mathbb{1}_{\{y_{pred}^{(i)} \neq y^{(i)}\}}$

↑ prediction value

$w^{(i)} = \begin{cases} 1 & \text{if } x^{(i)} \text{ is non-porn} \\ 10 & \text{if } x^{(i)} \text{ is porn} \end{cases}$

Orthogonalization for cat pictures: anti-porn

- 1. So far we've only discussed how to define a metric to evaluate classifiers. ← Place target
- 2. Worry separately about how to do well on this metric.

↑
Shot / aim at target.

$$\rightarrow J = \frac{1}{m} \sum_i^m w^{(i)} \ell(\hat{y}^{(i)}, y^{(i)})$$



Another example

Algorithm A: 3% error

✓ Algorithm B: 5% error ←

→ Dev/test



→ User images



If doing well on your metric + dev/test set does not correspond to doing well on your application, change your metric and/or dev/test set.