



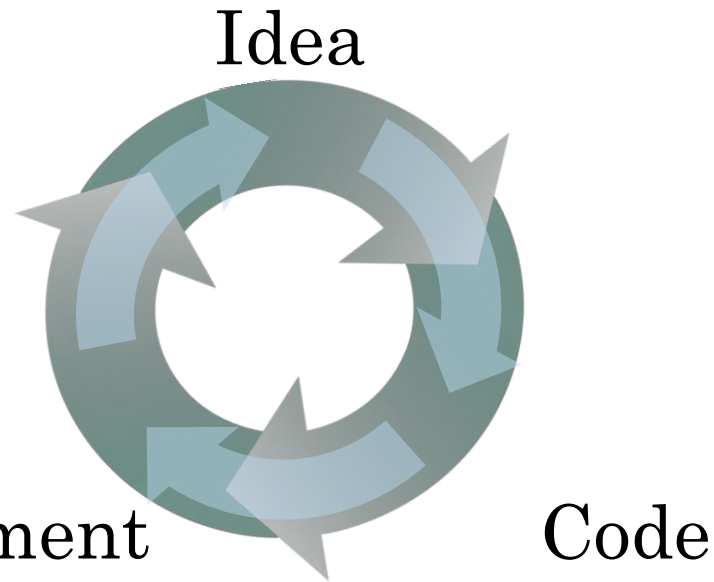
deeplearning.ai

Setting up
your goal

Single number
evaluation metric

Using a single number evaluation metric

use a single number to see
Model perf.



Classifier	Precision	Recall
A	95%	90%
B	98%	85%

of the examples your
classifier recognizes as
cats, what % are actually
cats

of All Images that are
cats in Actuality, what %
does your classifier
recognize as cats.

F₁ Score

92.4 ✓

91

- When precision ↑, Recall ↓
& vice versa,
which metric to consider?

- Instead, calculate the F1 score,
which is the "avg" of precision &
Recall

use F₁ as metric = $\frac{2}{\frac{1}{P} + \frac{1}{R}}$ → Harmonic
mean of P, R

Another example

Algorithm	US	China	India	Other
A	3%	7%	5%	9%
B	5%	6%	5%	10%
C	2%	3%	4%	5%
D	5%	8%	7%	2%
E	4%	5%	2%	4%
F	7%	11%	8%	12%

error by Algo A
on pictures of cats
submitted by USA users

Average

6

6.5

3.5 ✓ (lowest
Avg error)

5.25

3.75