Setting up
your goal

Train/dev/test
distributions

deeplearning.ai

Cat classification dev/test sets

development set, hold out cross validation set

Regions:

- US
- UK
- Other Europe
- South America
- India
- China
- Other Asia
- Australia

Dev

Test

Your dev and test set come from different distributions
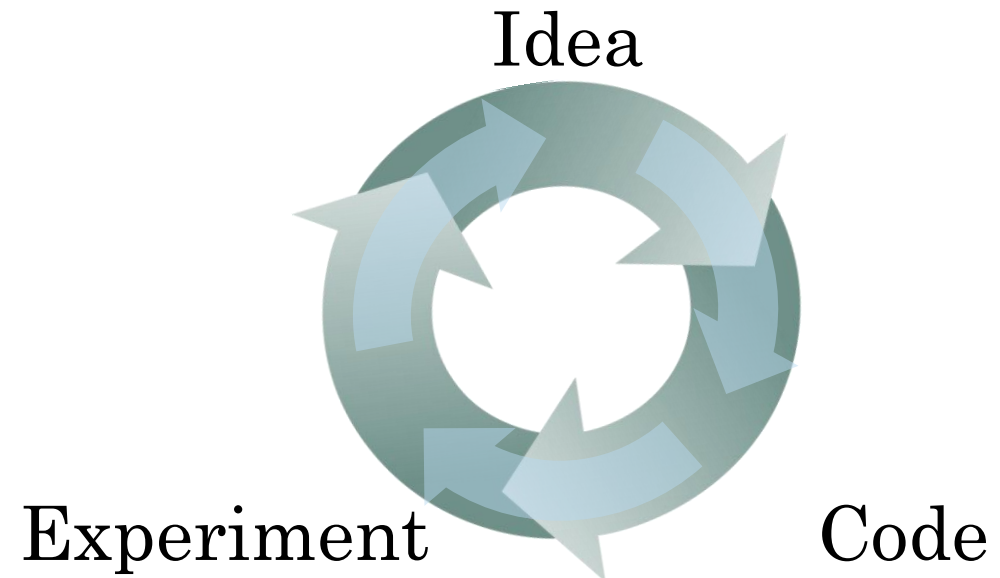
dev set + metric

Single real number evaluation metric

Randomly shuffle into dev/test

So that, both the dev and test sets have data from all eight regions and the dev and test sets really come from the same distribution, which is the distribution of all your data mixed together.

Idea

Experiment

Code

Andrew Ng

True story (details changed)

Optimizing on dev set on loan approvals for medium income zip codes

$x \longrightarrow y$ (repay loan?)
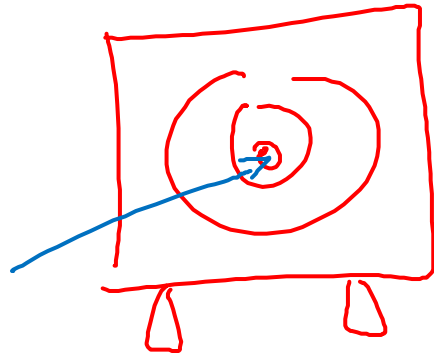
Tested on low income zip codes

~3 month

Andrew Ng
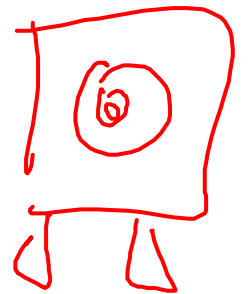
Guideline

Same   distribution

Choose a dev set and test set to reflect data you expect to get in the future and consider important to do well on.

Setting up the dev set, as well as the evaluation metric, is really defining what target you want to aim at. And hopefully, by setting the dev set and the training set to the same distribution, you're really aiming at whatever target you hope your machine learning team will hit. The way you choose your training set, will affect how well you can actually hit that target.

training

dev
metric

test

Andrew Ng