

# Error Analysis

Tuesday, September 1, 2020 9:24 PM

## ① Carrying out error analysis

- Look at dev examples to evaluate ideas

eg: if cat classifier has 90% accuracy (low error) and some dogs are mislabelled as cats

Should we try to make classifier do better on dog?

Error analysis:

- Get ~100 mislabeled dev set example
- Count how many are dogs

Case 1	Case 2
5 are dogs ↓ By fixing it we won't gain much error can go down 10% → 9.5%.	50 are dogs ↓ By fixing it we can potentially reduce error 50% → 5%.

- Evaluate multiple ideas in parallel

- eg:
- Fix pictures of dogs being recognised as cats
  - Fix great cats (tiger, panthers, etc.) being misrecognised
  - Improve performance on blurry images

Image	Dog	Great Cats	Blurry	Instagram	Comments
1	✓				Pitbull
2			✓	✓	
3			✓	✓	Rainy day at zoo
:	:	:		:	
% of total	8%	43%	61%	12%	

- Error analysis helps to make much better prioritization decisions, and understand how promising different approaches are to work on.

## ② Cleaning up incorrectly labelled data

- DL algorithms are quite robust to random errors in the training set (as long as dataset is big enough and the percentage of errors is low)
- They are less robust to systematic errors.
- In Dev/Test set, Incorrectly labeled can be added to the error analysis table
- We can determine if it's worth fixing incorrect labels through error analysis

Error analysis

Image	Dog	Great Cat	Blurry	Incorrectly labeled	Comments
...					
98				✓	Labeler missed cat in background
99		✓			
100				✓	Drawing of a cat; Not a real cat.
% of total	8%	43%	61%	6%	

Overall dev set error 10%

Errors due incorrect labels 0.6%

Errors due to other causes 9.4%

↑ ↗ not worth fixing      ↗ right be worth Fixing

Goal of dev set is to help you select between two classifiers A & B.

Andrew Ng

- Correcting incorrect dev/test set examples
  - Apply same process to your dev and test sets to make sure they continue to come from same distribution
  - Consider examining examples that are right as well as ones that are wrong
  - Train and dev/test data may now come from slightly different distributions (But it's okay!)

## ③ Build your first system quickly, then iterate

- Set up dev/test metric
- Build initial system quickly
  - ... performance analysis & review

- Build initial system quickly
- Use bias/variance analysis & error analysis to prioritize next steps  
(except if there is a lot of literature to build upon)  
to first model)