



deeplearning.ai

Error Analysis

Cleaning up
Incorrectly labeled
data

Incorrectly labeled examples

x



y

1

0

1

1

0

Training set.

DL algorithms are quite robust to random errors in the training set.

Systematic errors

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	<u>8%</u>	<u>43%</u>	<u>61%</u>	<u>6%</u>	

Overall dev set error 100%

Errors due incorrect labels 0.6% ←

Errors due to other causes 9.4% ←

2% ←

0.6% ←

1.4% ←

2.1%

1.9%

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

Correcting incorrect dev/test set examples

- Apply same process to your dev and test sets to make sure they continue to come from the same distribution
- Consider examining examples your algorithm got right as well as ones it got wrong.
(8.6%) *(20%)*
- Train and dev/test data may now come from slightly different distributions.