Check which percentage of each class presents error. Then, if some class is very present, optimize. If not, ommit.

Cleaning up Incorrectly labeled data

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

For dev set, check

Overall dev set error	10%	2%
Errors due incorrect labels	0.6% ←	0.6%
Errors due to other causes	9.4% ←	1.4%

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. (98.4) (2%)
- Train and dev/test data may now come from slightly different distributions.

Dev and test sets MUST come from the same distribution.