



deeplearning.ai

Setting up
your goal

Size of dev
and test sets

Old way of splitting data

→ 70, 30
Train Test

OR

60%, 20%, 20%
Train Dev Test

old method of splitting
works when data
size(total) < 10K
samples

Say data = 1 Million examples

→ 98% Train

→ 1% Test (10 K)

→ 1% Dev (10 K)

For large
data problem

Size of dev set

Set your dev set to be big enough to detect differences in algorithm/models you're trying out.

Size of test set

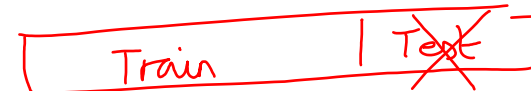
Set your test set to be big enough to give high confidence in the overall performance of your system.

Some times, you may not care about a high confidence in the performance of your system

In such cases, we omit the test set
⇒ Train, dev set

people have been often doing this when evaluating models, however they call the dev set → the "test" set

Not Recommended
(fine as long as model doesn't overfit dev set)



↳ It's actually the "dev" set
That you use to tune your model