

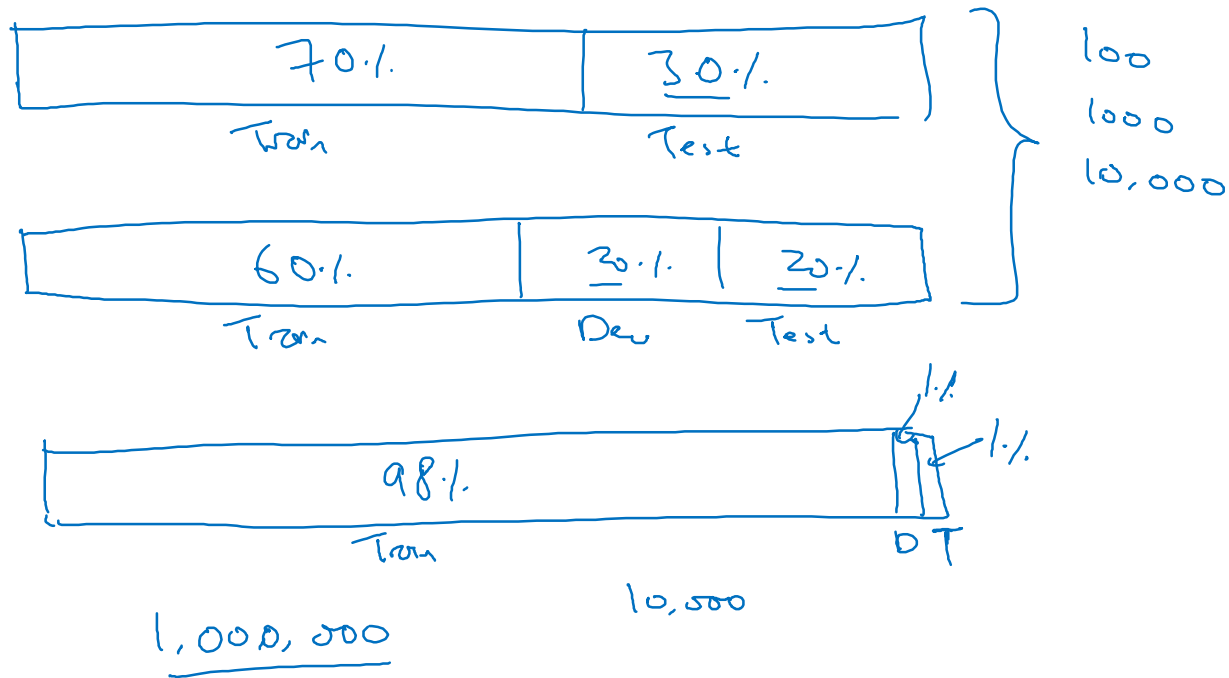


deeplearning.ai

Setting up
your goal

Size of dev
and test sets

Old way of splitting data



Size of dev set

A B

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

100 : small
└ 1%

1,000

10,000

100,000

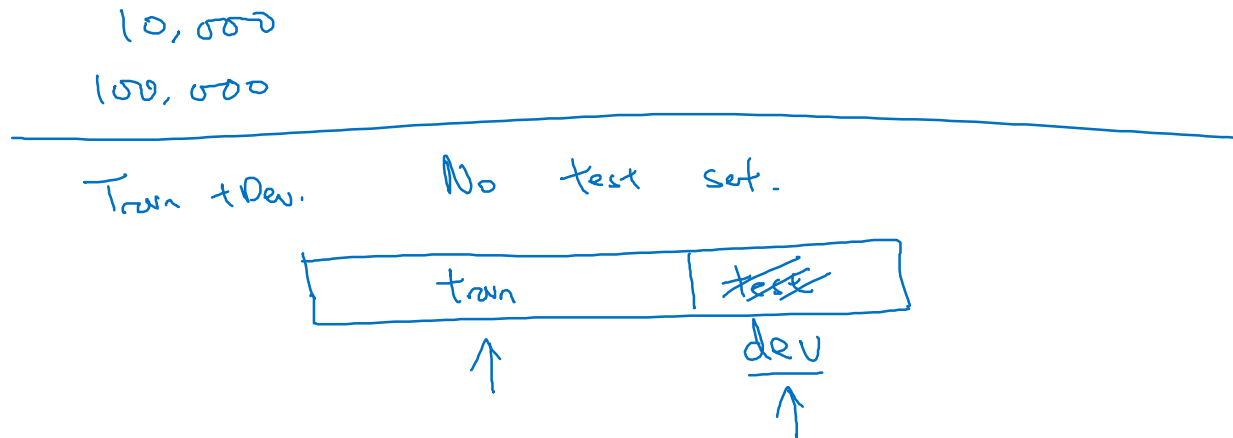
^A 97% → ^B 97.1%
0.1%
└

0.01%
└
0.001%

On the advertising

Size of test set

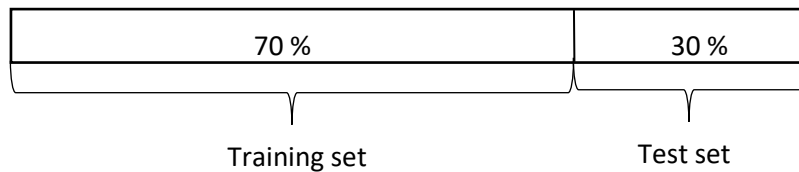
- Set your test set to be big enough to give high confidence in the overall performance of your system.



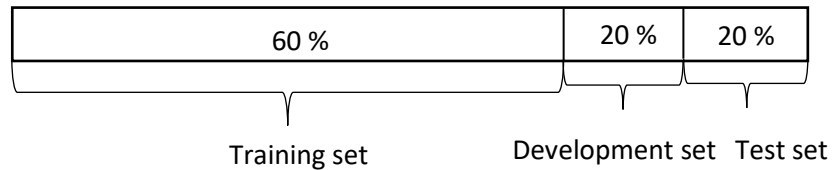
Size of the development and test sets

Old way of splitting data

We had smaller data set therefore we had to use a greater percentage of data to develop and test ideas and models.

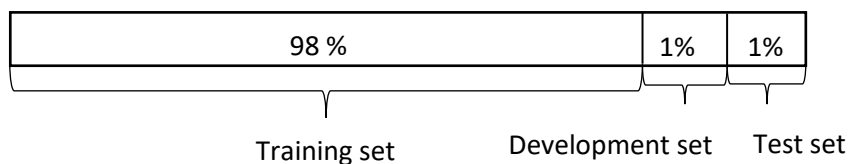


Or



Modern era – Big data

Now, because a large amount of data is available, we don't have to compromise as much and can use a greater portion to train the model.



Guidelines

- Set up the size of the test set to give a high confidence in the overall performance of the system.
- Test set helps evaluate the performance of the final classifier which could be less 30% of the whole data set.
- The development set has to be big enough to evaluate different ideas.