



deeplearning.ai

Mismatched training
and dev/test data

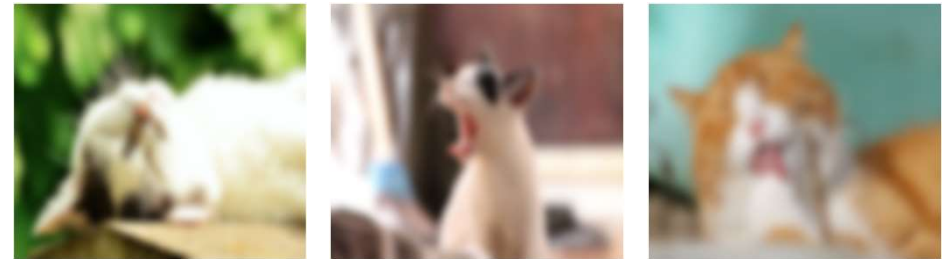
Training and testing
on different
distributions

Cat app example

Data from webpages



care about this
Data from mobile app

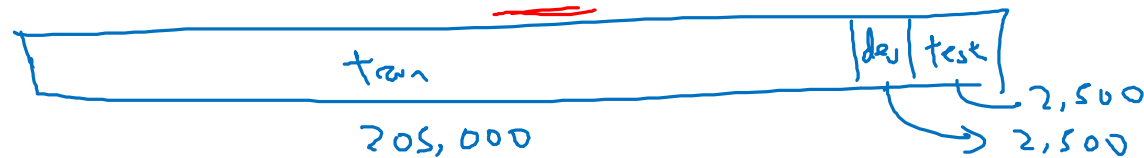


→ ≈ 200,000

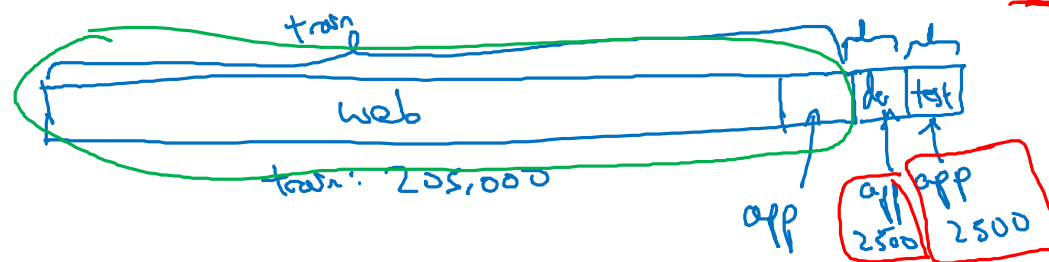
210,000
↓ shuffle

→ ≈ 10,000

~~Option 1:~~



Option 2:



$\frac{200K}{210K}$



2381 - web
119 - mobile app



Speech recognition example

Speech activated rearview mirror



Training

Purchased data

$\downarrow \downarrow$
 x, y

Smart speaker control

Voice keyboard

...

500,000 utterances

Dev/test

Speech activated
rearview mirror

