Mismatched training and dev/test data

Bias and Variance with mismatched data distributions
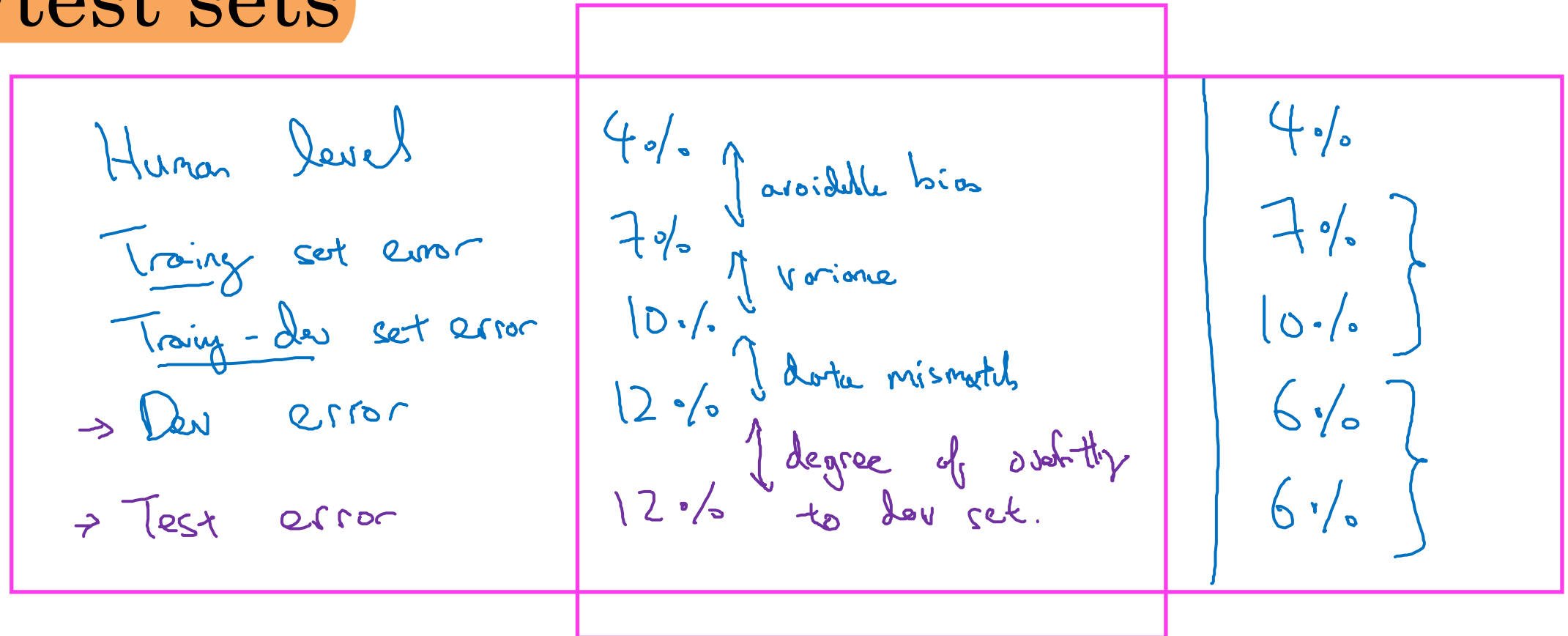
deeplearning.ai

Cat classifier example

Assume humans get ≈ 0% error.

Training error 1%
Dev error .... 10%   ↓ 9%

Training-dev set: Same distribution as training set, but not used for training

→ train   →train-dev   →dev test

NN saw this

Traing error	1% ↑	1% ↑
→ Traing-dev error	9%	↓ Variance 1.5%
→ Dev error	10%	10% ↓ data mismatch

Variance

Human error ---	0% ↑ Avoidable bias	
Traing error	10% ↓	10% ↑ Avoidable bias
Traing-dev error	11%	11% ↓ Variance
Dev error	12%	20% ↑ Data mismatch

Bias

Bias + Data mismatch

Andrew Ng
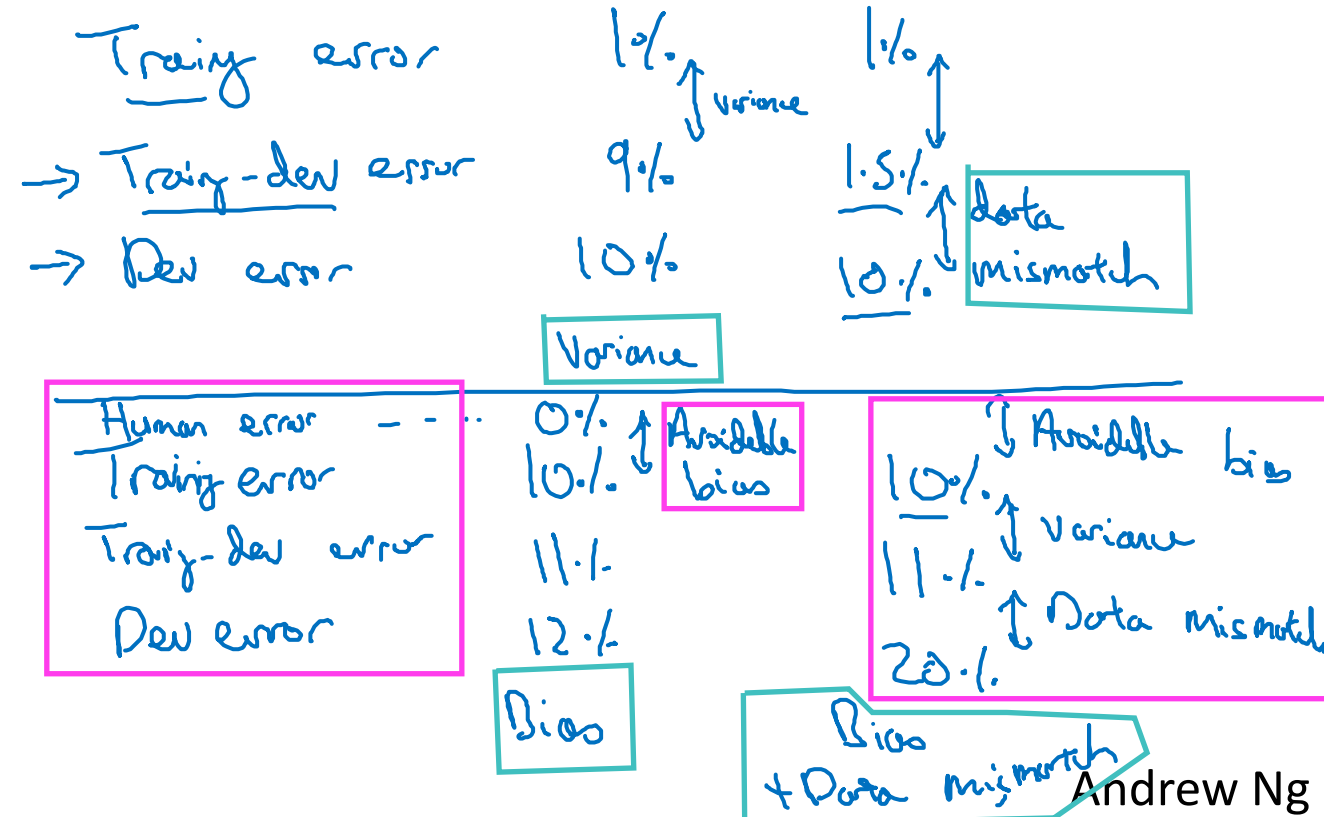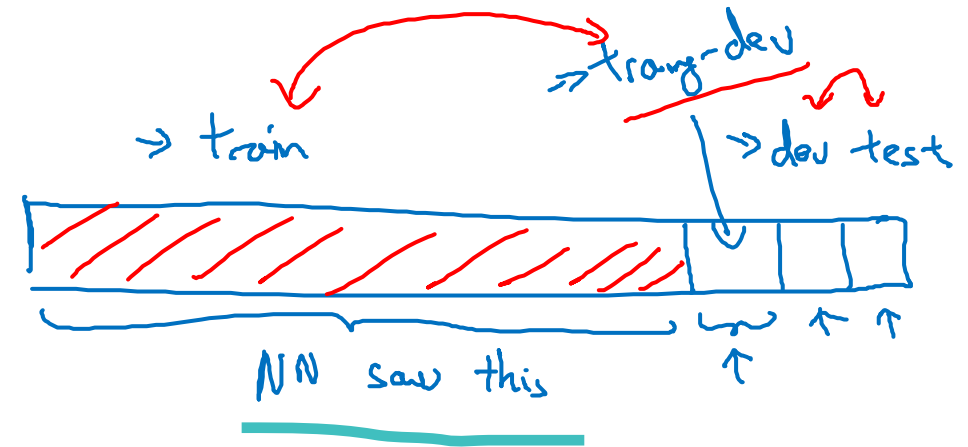
Bias/variance on mismatched training and dev/test sets

Human level

Traing set error

Traing - dev set error

→ Dev error

→ Test error

4%
⎫ avoidable bias
7%
⎫ variance
10%
⎫ data mismatch
12%
⎫ degree of overfitting
12%   to dev set.

4%

7% ⎫

10% ⎫

6% ⎫

6% ⎫

Andrew Ng

More general formulation

Rearview Mirror

	General speech recognition	Rearview mirror speech data.
Human level	"Human level" 4%	6%
Error on examples trained on	"Training error" 7%	6%
Error on examples not trained on	"Training-dev error" 10%	Dev/Test error 6%

Avoidable bias

Variance

Data mismatch