Mismatched training and dev/test data

Bias and Variance with mismatched data distributions

Cat classifier example

(training set and dev. test set from different distributions)
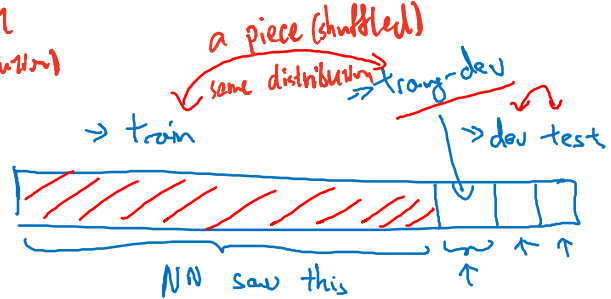
Assume humans get ≈ 0% error.

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

2 effects { variance
             data mismatch

define a new piece of data
Training-dev set: Same distribution as training set, but not used for training

a piece (shuffled)
← same distribution → Trang-dev
→ train   → dev test
NN saw this

Traing error	1% ↕ variance	1% ↕
→ Traing-dev error	9%	1.5% ↕ data mismatch
→ Dev error	10%	10%

Variance

Human error ----	0% ↕ Avoidable bias	↕ Avoidable bias
Traing error	10%	10% ↕ Variance
Traing-dev error	11%	11%
Dev error	12%	20% ↕ Data mismatch
	Bias	Bias + Data mismatch

Andrew Ng

Bias/variance on mismatched training and dev/test sets

Human level   4% ⎱ avoidable bias
Training set error   7% ⎱ variance
Training - dev set error   10% ⎱ data mismatch
→ Dev error   12% ⎱ degree of overfitting
→ Test error   12%   to dev set.

generally ↗

4%
7%   ⎱ hard
10%
6%   ⎱ easy
6%

funny example

More general formulation

give more additional insights

Reasview Mirror

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

Andrew Ng