

# Orthogonalization for cat pictures: anti-porn

- 1. So far we've only discussed how to define a metric to evaluate classifiers. ← Place target 
- 2. Worry separately about how to do well on this metric. 

An (shot at target)

$$\rightarrow J = \frac{1}{\sum w^{(i)}} \sum_{i=1}^m w^{(i)} \ell(\hat{y}^{(i)}, y^{(i)})$$



# Another example

Algorithm A: 3% error

✓ Algorithm B: 5% error ↙

→ Dev/test



→ User images



If doing well on your metric + dev/test set does not correspond to doing well on your application, change your metric and/or dev/test set.



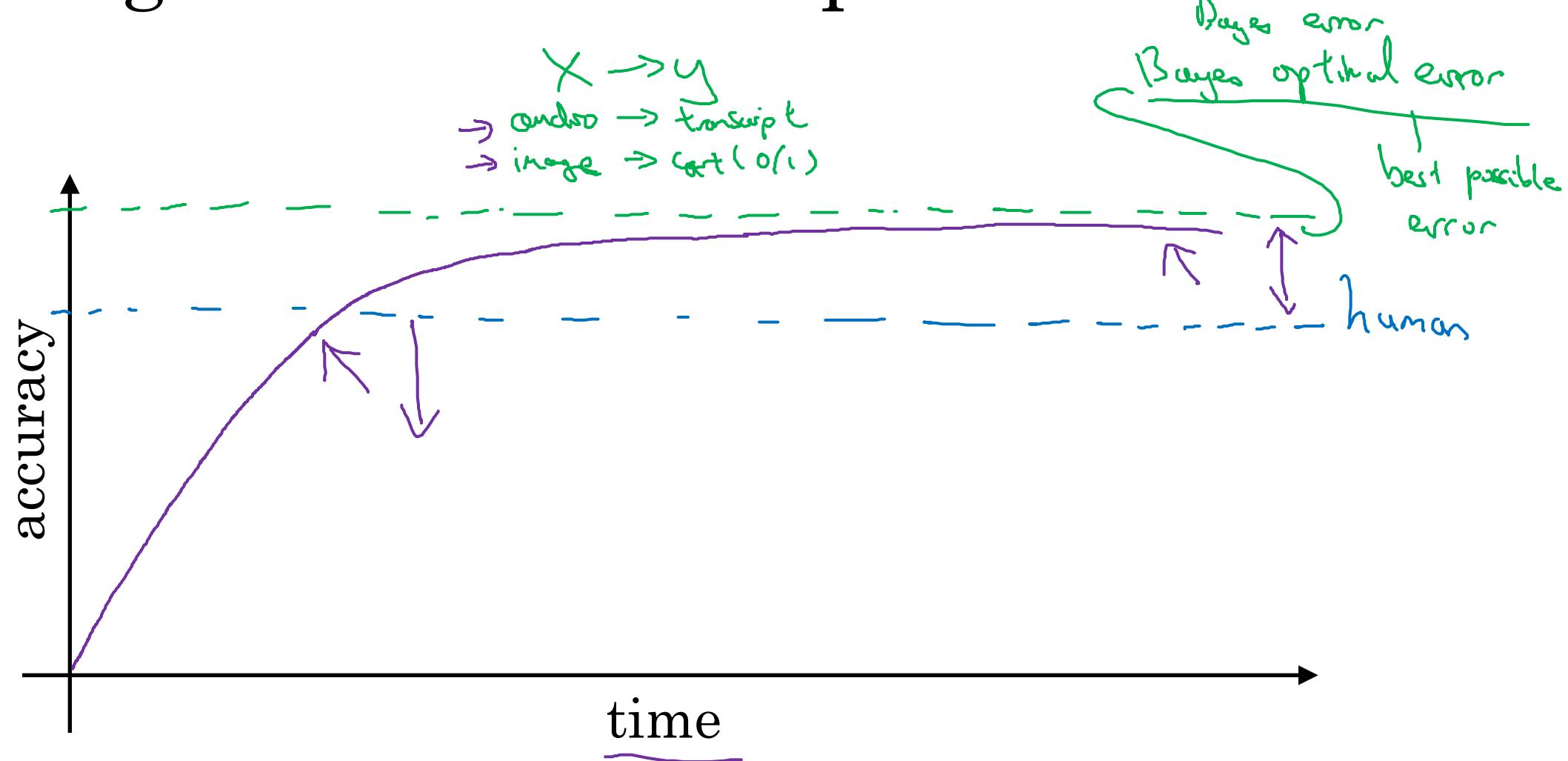
deeplearning.ai

Comparing to human-level performance

---

Why human-level performance?

# Comparing to human-level performance



# Why compare to human-level performance

Humans are quite good at a lot of tasks. So long as ML is worse than humans, you can:

- - Get labeled data from humans.  $(x, y)$
- - Gain insight from manual error analysis:  
Why did a person get this right?
- - Better analysis of bias/variance.



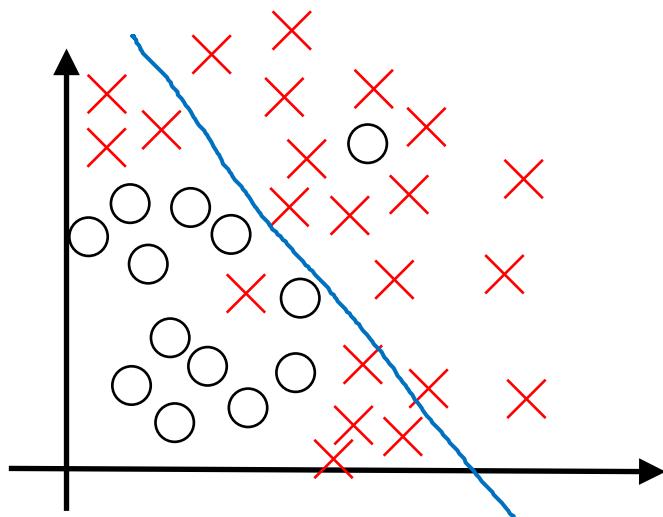
deeplearning.ai

Comparing to human-level performance

---

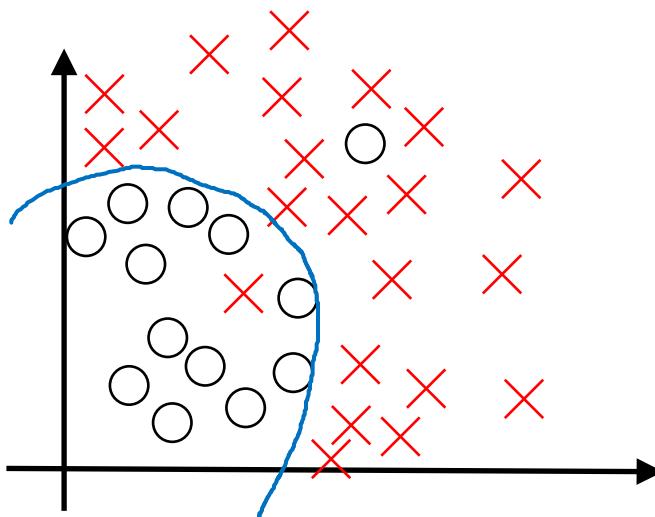
Avoidable bias

# Bias and Variance

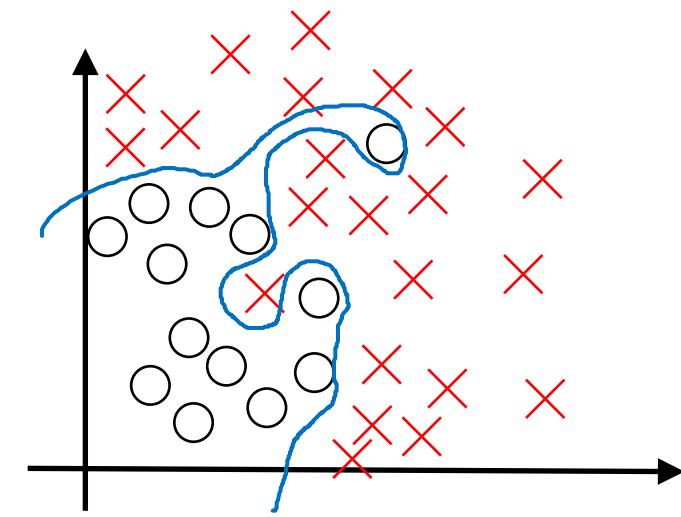


high bias

*underfitting*



"just right"



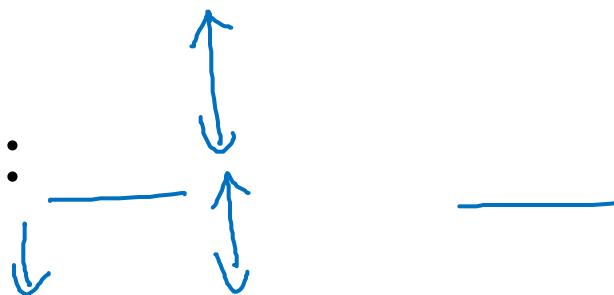
high variance

*overfitting*

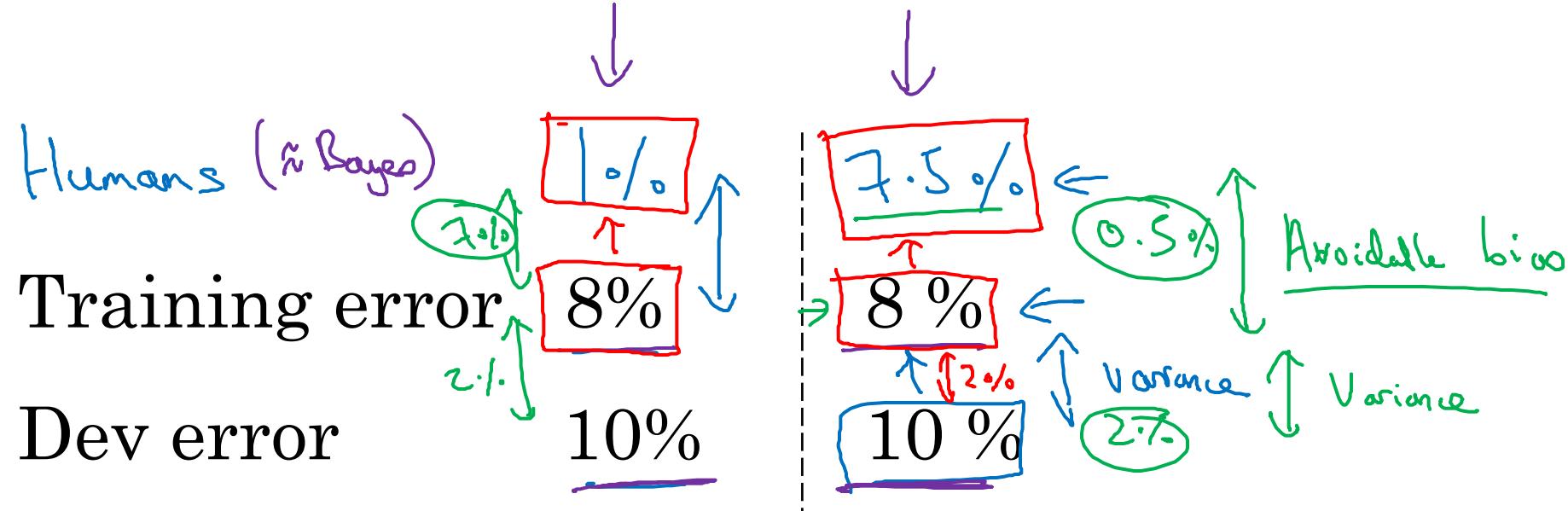
# Bias and Variance

Cat classification

Human-level  $\approx 0\%$



# Cat classification example



Focus on

bias

Focus on

Variance

Human-level error as a proxy for Bayes error.



deeplearning.ai

Comparing to human-level performance

---

Understanding  
human-level  
performance

# Human-level error as a proxy for Bayes error

Medical image classification example:

Suppose:

- (a) Typical human ..... 3 % error
- (b) Typical doctor ..... 1 % error
- (c) Experienced doctor ..... 0.7 % error
- (d) Team of experienced doctors .. 0.5 % error



What is “human-level” error?

$$\text{Baye error} \leq \underline{0.5\%}$$

# Error analysis example

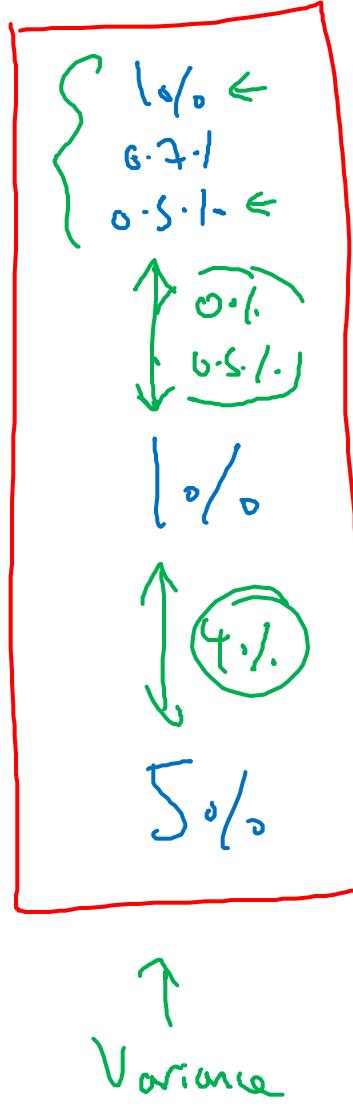
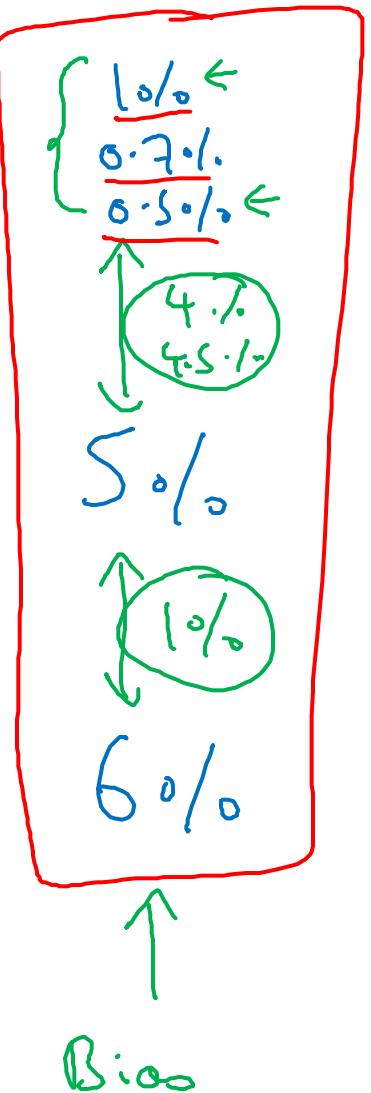
Human (proxy for Bayes error)



Training error

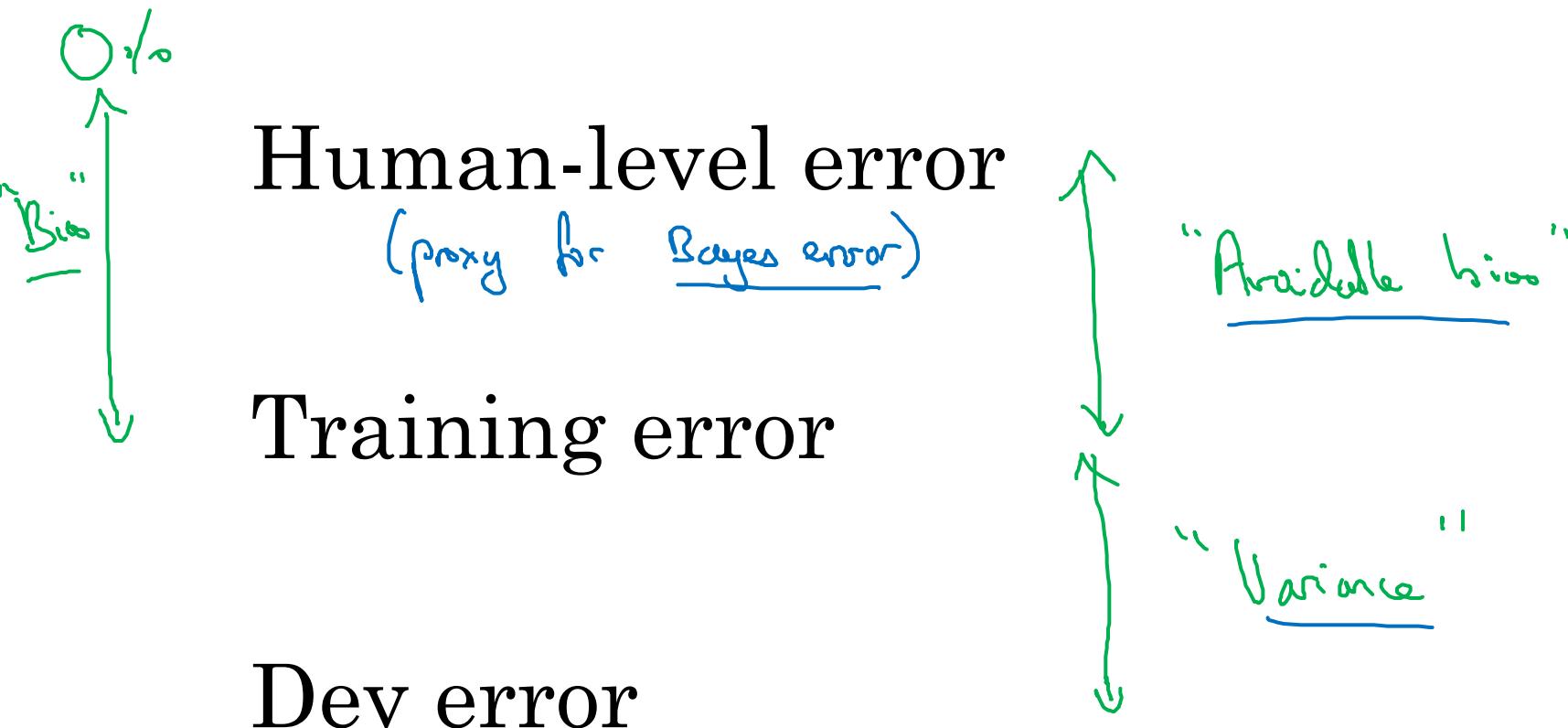


Dev error



$$\begin{aligned} &\rightarrow \frac{0.7\%}{0.5\%} = 1.4\% \\ &\rightarrow 0.2\% \\ &\rightarrow 0.0\% \\ &\rightarrow 0.7\% \\ &\rightarrow 0.1\% \\ &\rightarrow 0.8\% \end{aligned}$$

# Summary of bias/variance with human-level performance





deeplearning.ai

Comparing to human-level performance

---

Surpassing human-level performance

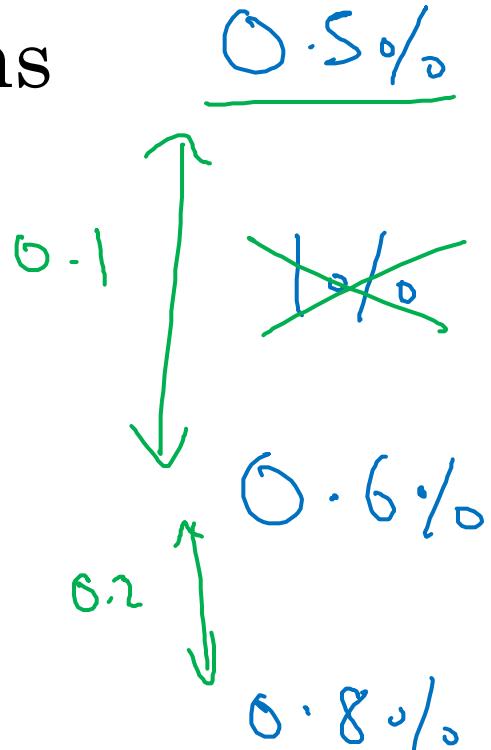
# Surpassing human-level performance

Team of humans

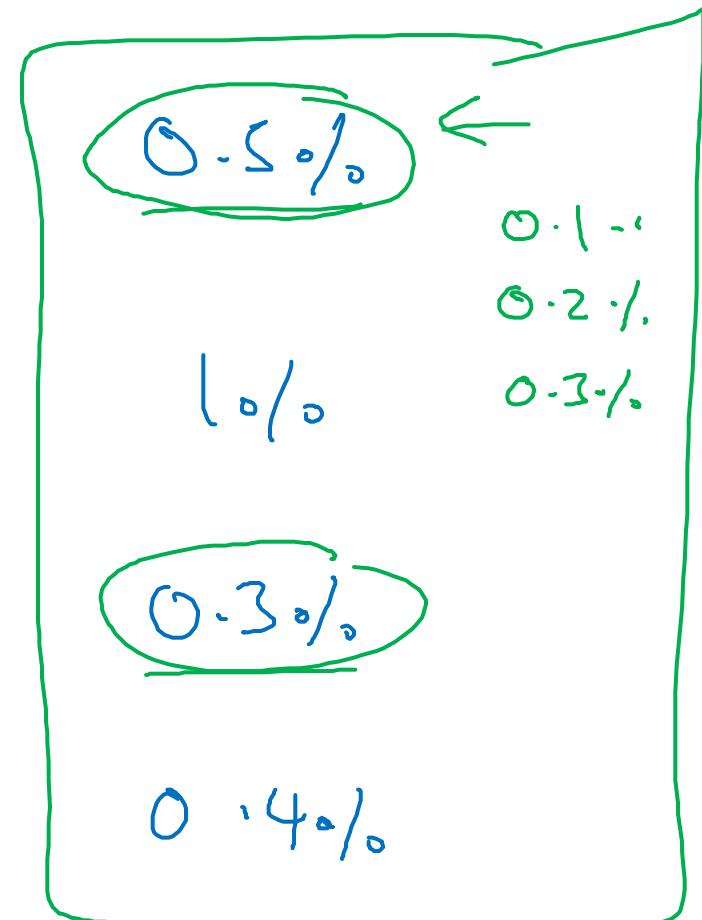
One human

Training error

Dev error

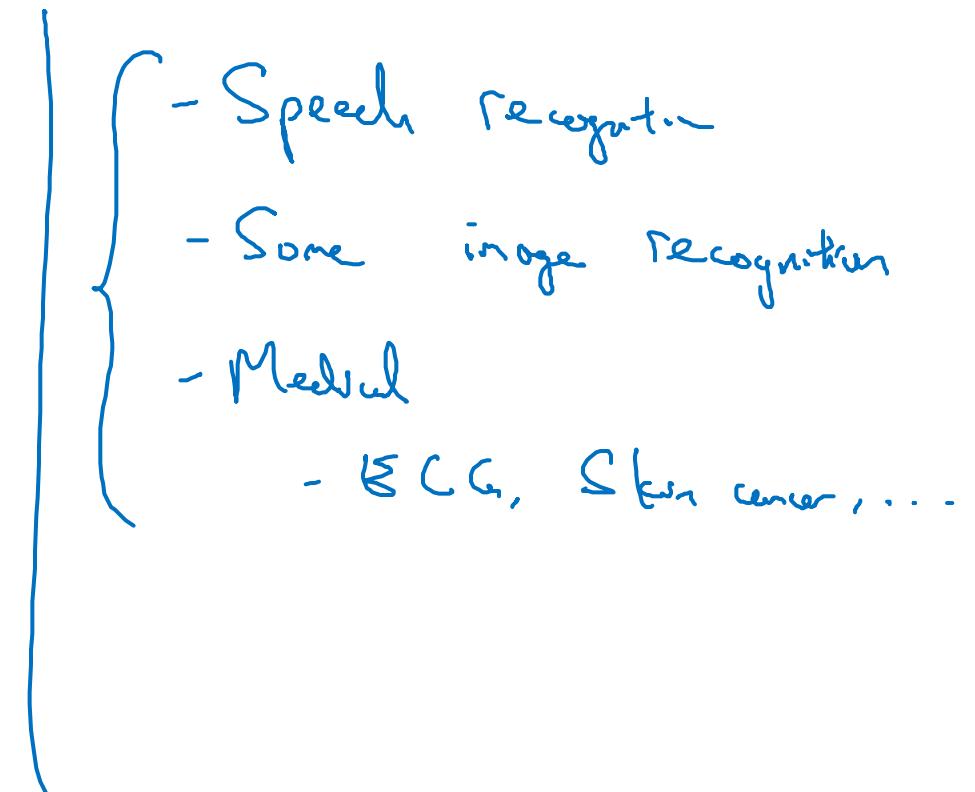


What is avoidable bias?



# Problems where ML significantly surpasses human-level performance

- - Online advertising
- - Product recommendations
- - Logistics (predicting transit time)
- - Loan approvals



Structural data

Not natural perception

Lots of data



deeplearning.ai

Comparing to human-level performance

---

Improving your model performance

# The two fundamental assumptions of supervised learning

1. You can fit the training set pretty well.



$\sim$  Avoidable bias

2. The training set performance generalizes pretty well to the dev/test set.



$\sim$  Variance

# Reducing (avoidable) bias and variance

