deeplearning.ai

Comparing to human-level performance

Understanding human-level performance

Human-level error as a proxy for Bayes error

Medical image classification example:

Make a diagnosis classification

Suppose:

(a) Typical human ..... 3 % error

(b) Typical doctor ..... 1 % error

(c) Experienced doctor ..... 0.7 % error

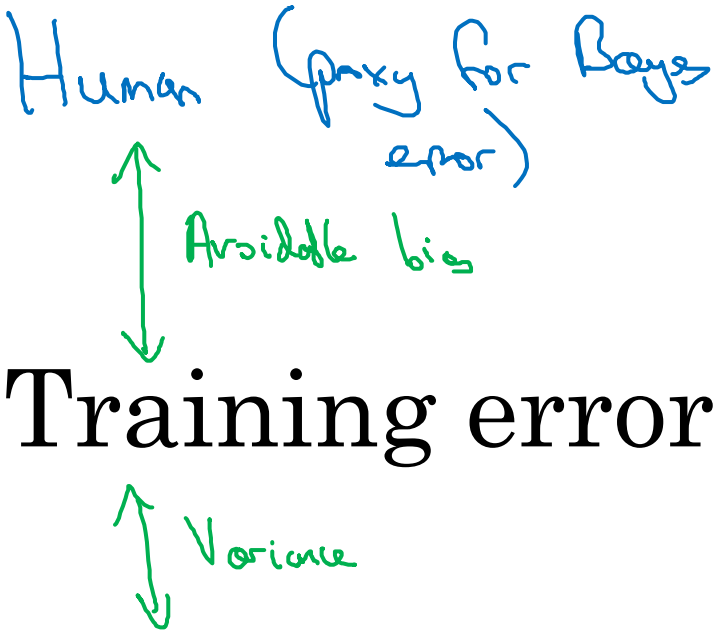
(d) Team of experienced doctors .. 0.5 % error

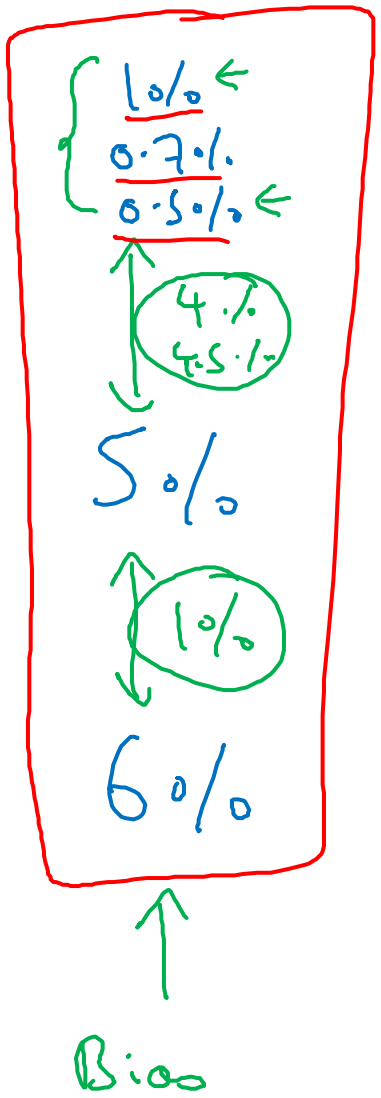
Bayes error $\leq$ 0.5%

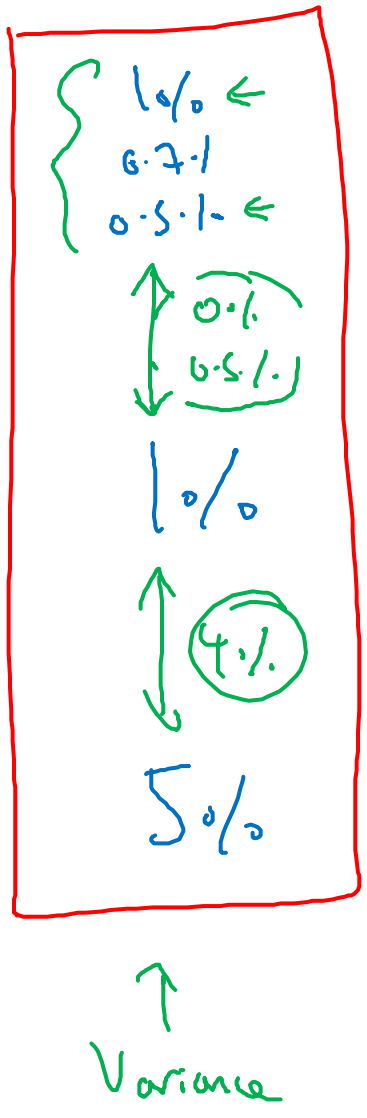
What is "human-level" error?

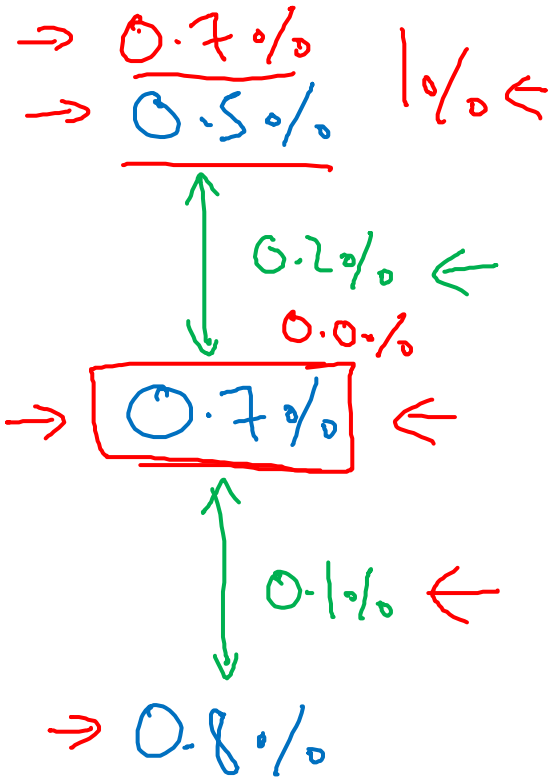
One of the most useful ways to think of human error is as a proxy or an estimate for Bayes error.

Error analysis example









Summary of bias/variance with human-level performance

Human-level error

(proxy for Bayes error)

approximation

"Avoidable bias"

Training error

"Variance"

Dev error

whether your algorithm's able to generalize from the training set to the dev set.