deep Learning_c3_week1

Created:	19/02/13 18:18	**Updated:**	19/02/13 20:51

orthogonalization

- each control does a specific task and doesn't affect other controls.

single number evaluation metric

- confusion matrix: regression/recall
- F1 scores: `F1 = 2 / ((1/P) + (1/R))`

satisfying and optimizing metric

- facing many metrics, hard to judge which model is better.
    - e.g.

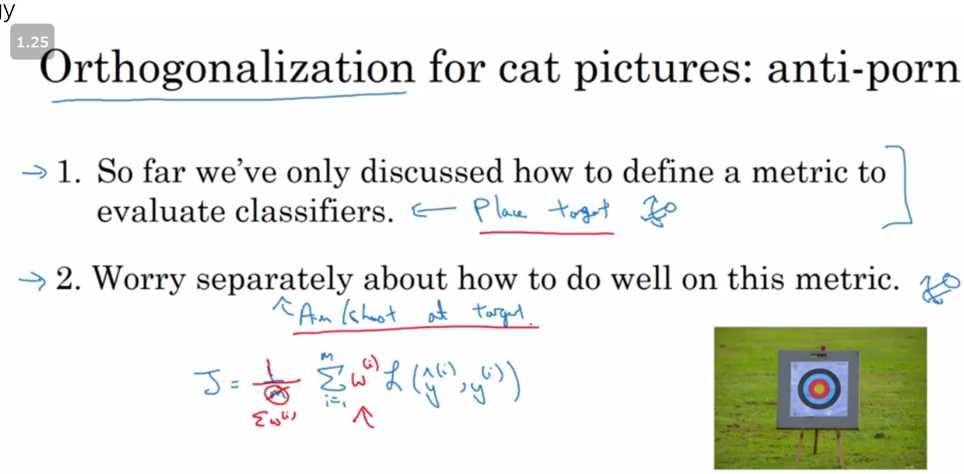
Classifier	F1	Running time
A	90%	80 ms
B	92%	95 ms
C	92%	1,500 ms

- choosing a single **optimizing** metric and decide that other metrics are **satisfying**.
    - `Maximize F1 # optimizing metric`
    - `subject to running time < 100ms # satisficing metric`

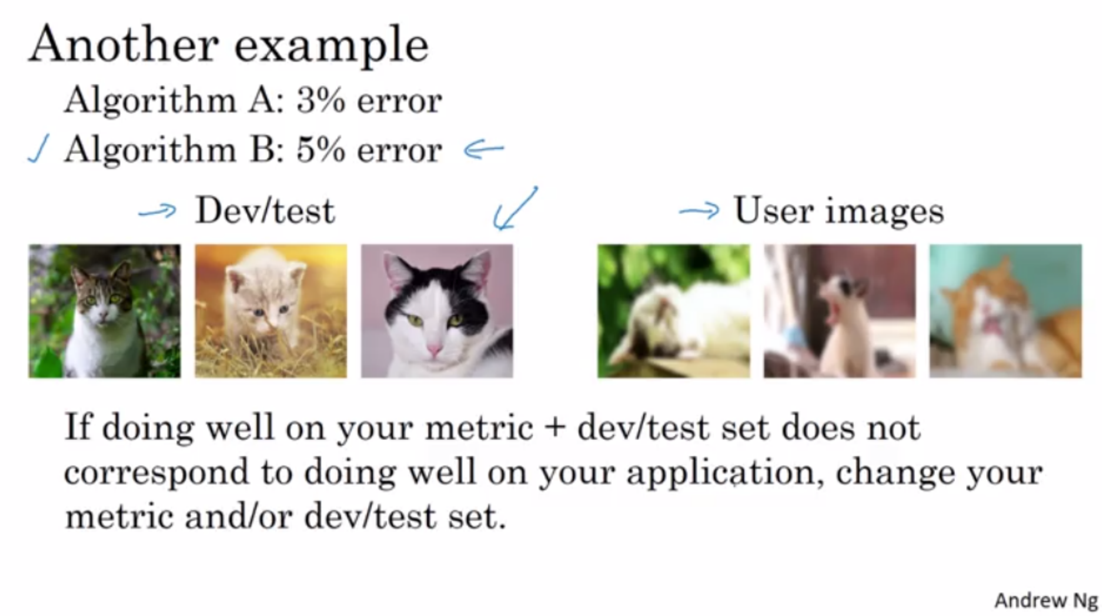
train/dev/test sets distribution

- come from same distribution
- Choose dev set and test set to reflect data you expect to get **in the future** and consider important to do well on

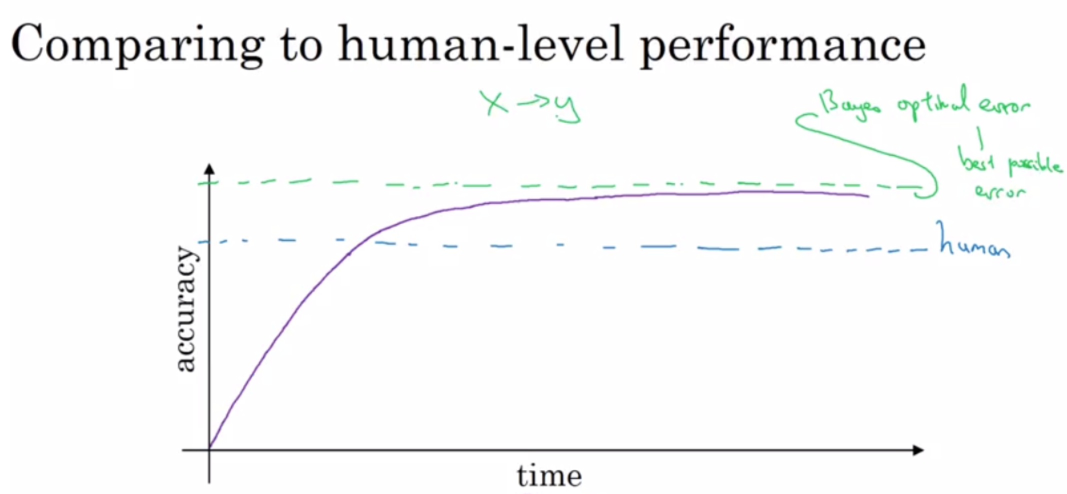
when to change dev/test sets and metrics

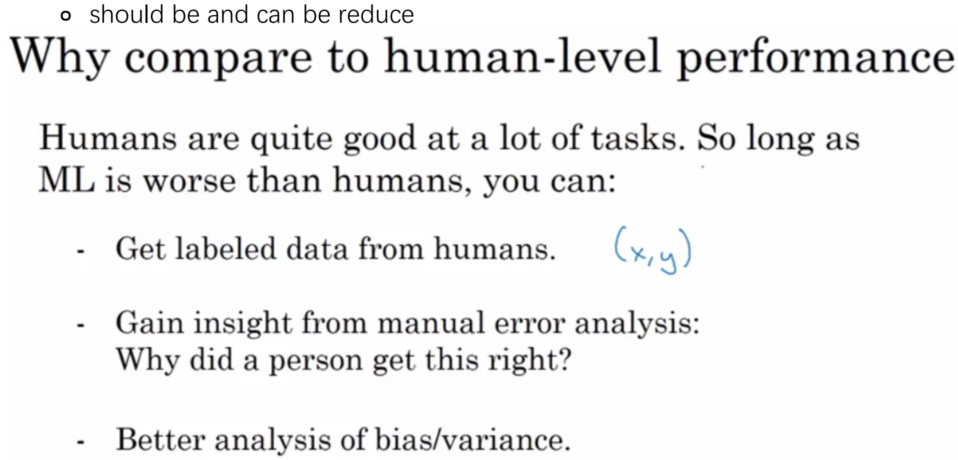
- analogy
    - 

        - Figure out how to define a metric that captures what you want to do - place the target.
        - Worry about how to actually do well on this metric - **how to aim/shoot accurately at the target.**
- so

- 

human-level performance

- human-level error can be used to estimate **bayes error**
- 

- avoidable error
  - define as `Avoidable bias = Training error - Human (Bayes) error`
  - should be and can be reduce
- 

Improve model performance

- fundamental assumptions of supervised learning:
    i. fit the training set pretty well (achieve low **avoidable bias**.)
    ii. The training set performance generalizes pretty well to the dev/test set. (**variance** is not too bad.)
- To improve your deep learning supervised system
    i. Look at the difference between human level error and the training error - **avoidable bias**.
    ii. Look at the difference between the dev/test set and training set error - **Variance**.
    iii. If **avoidable bias** is large you have these options:
        - Train bigger model.
        - Train longer/better optimization algorithm (like Momentum, RMSprop, Adam).
        - Find better NN architecture/hyperparameters search.
    iv. If **variance** is large you have these options:
        - Get more training data.
        - Regularization (L2, Dropout, data augmentation).
        - Find better NN architecture/hyperparameters search.