

The Machine Learning Process

Machine learning problems involve multiple steps

1) Data collection

- camera
- database



2) Preprocessing

- common size, center
- reject poor quality images

Time consuming!

Important!

3) Feature extraction

What attributes are relevant?

- Collect measurements in a vector

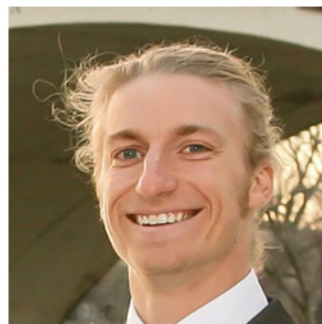
4) Collect training data



Barry



Gauss



Dave

Labels



5) Apply a model - captures salient aspects³
of the data

- face identity represented by differences in landmarks relative to the mean
- movie ratings are a weighted combination of nominal taste profiles



"All models are wrong, but some are
useful" George E. P. Box
(1919 - 2013)

Two general categories of problems:

4

1) Supervised learning

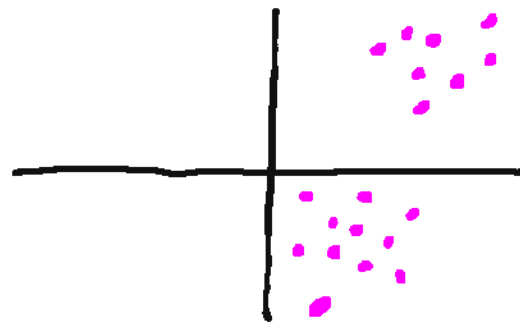
- examples with correct outcomes
- train model with examples
- apply trained model to new data

Ex: Classification  VS 

2) Unsupervised learning

- no training data/examples

Ex: clustering



Copyright 2019
Barry Van Veen