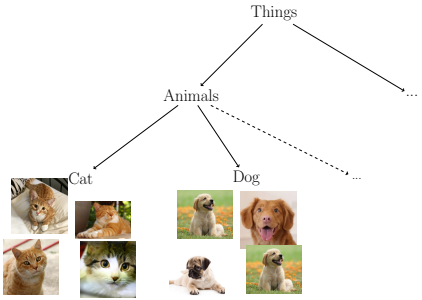
Biological Vision and Applications
Module 03-08: Taxonomy Learning

Hiranmay Ghosh

Taxonomy

Organizing concepts in a hierarchy

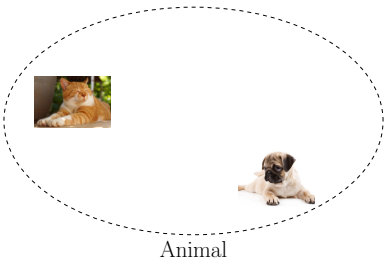


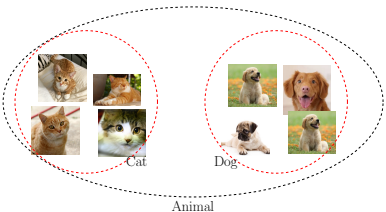
- Learned top-down, or bottom-up ?

Taxonomy Learning

The cognitive science viewpoint

- Psychologists suggest that it is learned top-down with experience





Taxonomy is a tradeoff

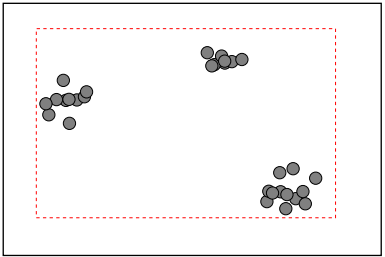
... between complexity of hypothesis and it's goodness of fit (with data)

- Complexity of hypothesis
  - Human mind accepts the simplest theory
  - A hypothesis with one class is simplest
    - ... more classes $\rightarrow$ more complexity
- "Goodness of fit"
  - Probability of data being explained by the hypothesis
  - Tighter the enclosed area, better is the goodness of fit

Example

Hypothesis 1

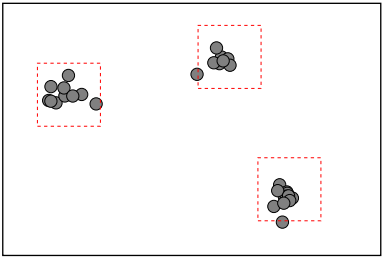
Simple (1 class)
Poor goodness of fit (large bounding box)



Feature space

Hypothesis 2

Complex (3 classes)
Better goodness of fit (small bounding boxes)



Feature space

- Which one is accepted ?

Bayesian approach for hypothesis selection

- Posterior probability of a hypothesis: $P(h_i \mid d) = k_1.P(h_i).P(d \mid h_i)$
  - Choose $h^* = \text{argmax}_i P(h_i \mid d)$
  - We can ignore k_1 (for comparisons)
- Prior for hypothesis: $P(h_i) = f(c_i)$
  - c_i: number of classes in hypothesis h_i (complexity)
  - $f(c_i)$: a exponentially decreasing function on c_i
    - $P(h_i) = k_2^{-c_i}$

Bayesian approach for hypothesis selection

... contd.

- "Goodness of fit": $P(d \mid h_i) = \left(\frac{k_3}{A}\right)^n$
  - Inverse of probability of the data to fit in the designated area (in feature space)
  - n: Total number of data points
  - A: Total area for all the category spaces
  - k_3: A constant (scaling factor for A)
- Posterior: $P(h \mid d) = k_1.k_2^{-c_i}.\left(\frac{k_3}{A}\right)^n$
- Generally ...
  - Simpler hypothesis dominates with sparse data
    - When we have seen less number of instances, we tend to believe that they are all in the same group
  - Goodness of fit dominates with dense data
    - When we have seen more instances, we tend to classify them in subgroups

Quiz

No quiz for module 03-08

End of Module 03-08