Valiant’s Theorem

Emmy Huang

Hume's Problem of Induction

Q: If you observe 500 black ravens, what basis do you have for supposing that the next one you observe will also be black?

Thoughts?

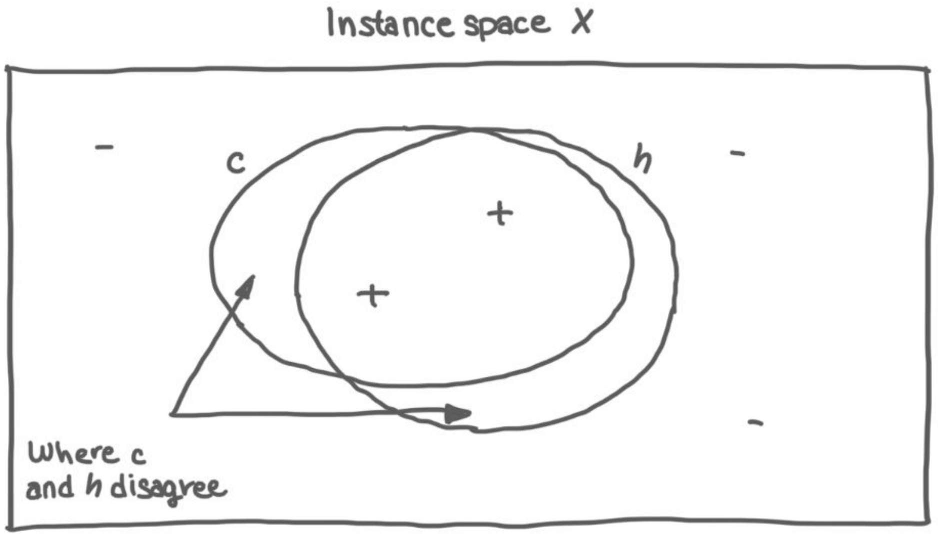
- Bayes' Theorem
  - Assumes all ravens are drawn from same distribution
- Computational Learning Theory
  - Learning does happen but how?
  - Not equal footing
  - Why does this work?

PAC-learning (Probably Approximately Correct)

- High probability → mostly correct predictions
- S: sample space
- f: concept
- C: concept class
- D: probability distribution
- Goal: given m examples x_i drawn independently from D, we know $f(x_i)$ → output hypothesis language h such that h disagrees with f no more than ε of the time

Equation and Visualization

- $error(h) = P(h(x) \neq f(x) \mid x \text{ drawn from } D) \leq \varepsilon$

- 

Valiant's Theorem

- In order for the output hypothesis h to agree with $1 - \varepsilon$ of the future data drawn from D with probability $1 - \delta$ over the choice of samples, it suffices to find any hypothesis h that agrees with:

$$m \geq \frac{1}{\varepsilon} \log\left(\frac{|C|}{\delta}\right)$$

samples chosen independently from D.

Proof

- Bad hypothesis h
  - Disagrees with f for at least ε fraction of data
- Thus: $\Pr[h(x_1) = f(x_1), \ldots, h(x_m) = f(x_m)] < (1-\varepsilon)^m$
- Probability that there exists a bad hypothesis h in C that agrees with sample data?
- $\Pr[$there exists a bad h that agrees with f for all samples$] < |C|(1-\varepsilon)^m$
- Set equal to δ and solve for m:
- $$m = \frac{1}{\varepsilon}\log\left(\frac{|C|}{\delta}\right)$$

Further Exploration

- Infinite concept classes? Rectangle in plane?
- Shattering, VC Dimension

Thank you!