Probability formulas

$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$ → Product Rule

$P(h|D) = \frac{P(D|h)P(h)}{P(D)}$ → Posterior probability

$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$ → Law of Total Probability

Brute Force Max a Posteriori + Concept Learning

Assumptions:

1. Training data is noise free.
2. Target concept c is in the H. (hyp. space)
3. Every hypothesis has equal priors

Then: $P(h) = \frac{1}{|H|}$ $P(D|h)$: probability of data D given hypothesis h is 1 if data is consistent with h, 0 otherwise.

$$P(h|D) = \frac{\overset{1 \text{ or } 0}{P(D|h)}\ \overset{\frac{1}{|H|}}{P(h)}}{P(D)} \rightarrow P(h|D) = \frac{1 \cdot \frac{1}{|H|}}{P(D)} = \frac{1 \cdot \frac{1}{\cancel{|H|}}}{\frac{|VS_{H,D}|}{\cancel{|H|}}}$$

$$= \frac{1}{|VS_{H,D}|}$$

$|VS_{H,D}|$ → Subset of hypotheses from H consistent with data.

$$P(D) = \frac{|VS_{H,D}|}{|H|}$$

$$P(h|D) = \begin{cases} \frac{1}{|VS_{H,D}|} & \text{if } h \text{ is consistent with } D \\ 0 & \text{otherwise} \end{cases}$$

Initially all $h = \frac{1}{|H|}$ as we add data posterior of inconsistent hypotheses become zero while consistent ones have a probability of $\frac{1}{\#\text{consistent ones}}$