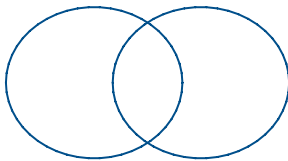


Probability

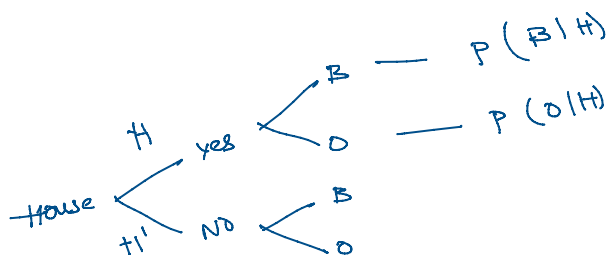
Friday, December 23, 2022 11:11 AM



$H \setminus B$
 $H \cap B$

$$\begin{aligned} P(B) &= 0.6 \\ P(O) &= 0.4 \\ P(H|B) &= 0.25 \\ P(H|O) &= 0.2 \end{aligned}$$

$$\begin{aligned} P(B|H) &= P(B \cap H) / P(H) \\ P(H) &= P(B \cap H) + P(H \cap O) \\ &= P(H|B) * P(B) + P(H|O) * P(O) \end{aligned}$$



$$\begin{aligned} P(H|B) &= \frac{P(H \cap B)}{P(B)} \Rightarrow P(H \cap B) = \frac{P(H|B) * P(B)}{1} \\ P(H|O) &= \frac{P(H \cap O)}{P(O)} \Rightarrow P(H \cap O) = \frac{P(H|O) * P(O)}{1} \end{aligned}$$

Spam Assassin works by having users train the system. It looks for patterns in the words in emails marked as spam by the user. For example, it may have learned that the word "free" appears in 30% of the mails marked as spam, i.e., $P(\text{Free} | \text{Spam}) = 0.30$. Assuming 1% of non-spam mail includes the word "free" and 50% of all mails received by the user are spam, find the probability that a mail is spam if the word "free" appears in it.

Draw the probability tree diagram

$$\begin{aligned} P(\text{Free} | \text{Spam}) &= 0.3 \\ P(\text{Free} | \text{Non-spam}) &= 0.01 \\ P(\text{Spam}) &= 0.5 \\ P(\text{Spam} | \text{Free}) &= ? \\ P(\text{NS}) &= 1 - P(S) = 1 - 0.5 = 0.5 \end{aligned}$$

