Bayes theoram:-

Bayes theoram describes the probability of an event based on prior knowledge of conditions that might be related to the event. If we know the conditional probability $P(A/B)$, we can use the bayes rule to find out the reverse probabilities $P(B/A)$.

Prior knowledge : P(Positive result shows / Person has Disease)

Bayes: P(Person has Disease / It shows Positive result)

Baysian Spam filter:-

A baysian spam filter uses information about previously seen email messages to guess whether an incoming email is spam. Baysian spam filter looks for a particuler word occurences in message. For a particuler word w, the probability that w apears in a spam email message is estimated by determining the number of times w apears in a message from a large set of messages known to be spam and

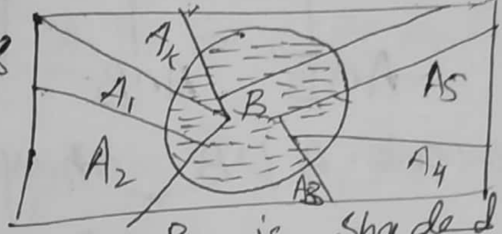
Bayes theoram :- If the events $A_1, A_2 \ldots A_k$ form a partition of Sample Space S, that is the events A_i are mutually exclusive and their union is S, and if B is any other event of S such that it can only occur if one of the A_i occurs then for any i

$$P(A_i/B) = \frac{P(A_i) \cdot P(B/A_i)}{\sum P(A_i) \cdot P(B/A_i)}, \quad \text{for } i = 1, 2 \ldots, k$$

By Multiplication rule

$$P(B \cap A_i) = P(B) P(A_i/B)$$

$$= P(A_i) \cdot P(B/A_i)$$



B is shaded

Equating

$$P(B) \cdot P(A_i/B) = P(A_i) \cdot P(B/A_i)$$

$$P(A_i/B) = \frac{P(A_i) \cdot P(B/A_i)}{P(B)} \quad \text{--- (1)}$$

we may write the event B as $B = S \cap B$

$$B = S \cap B$$

$$B = (A_1 \cup A_2 \cup A_3 \ldots \cup A_k) \cap B$$

$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup \ldots \cup (A_k \cap B)$$

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + \ldots + P(A_k \cap B)$$

$$= P(A_1) \cdot P(B/A_1) + P(A_2) \cdot P(B/A_2) + \ldots + P(A_k) \cdot P(B/A_k)$$

$$P(B) = \sum_{i=1}^{k} P(A_i) \cdot P(B/A_i)$$

This result is known as theoram on Total probability, put it in eq (1)

$$P(A_i/B) = \frac{P(A_i) \cdot P(B/A_i)}{\sum P(A_i) \cdot P(B/A_i)}$$

Hence the result.

Scanned with CamScanner

number of times in ~~non~~ non-spam.

Prior knowledge : $P\left(\text{word}/\text{spam}\right)$

Bayes : $P\left(\text{spam}/\text{word}\right)$

Question:-

Suppose that we have found that the word "Rolex" occurs in 250 of 2000 messages known to be spam and in 5 of 1000 messages known not to be spam. Estimate the probability that an incoming message containing the word "Rolex" is spam, assuming that it is equally likely that an incoming message is spam or not spam. If our threshold for rejecting a message as spam is 0.9, will we reject such messages?

Scanned with CamScanner

R: Word Rolex occurs

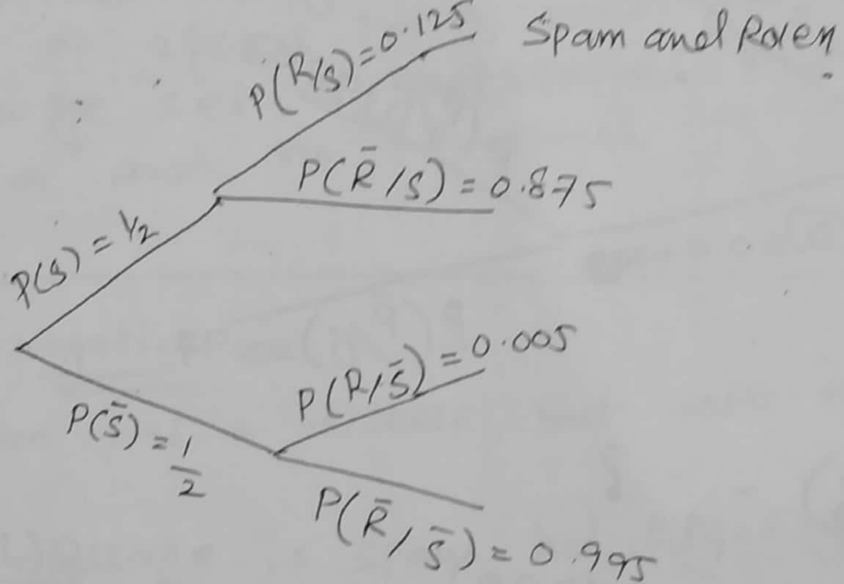
S: Message is spam

$\bar{S}$: Message is not spam

$\bar{R}$: Rolex not occurs

P(Message is spam / it contains the word Rolex) = ?

P(Spam) = P(Not Spam) = $\frac{1}{2}$



$$P(S/R) = \frac{P(R/S) \cdot P(S)}{P(R/S) \cdot P(S) + P(R/\bar{S}) \cdot P(\bar{S})}$$

$$= \frac{P(S \cap R)}{P(S \cap R) + P(\bar{S} \cap R)} = \frac{P(R \cap S)}{P(R)}$$

$$= \frac{0.125 \times 0.5}{0.125 \times 0.5 + 0.005 \times 0.5}$$

$$= 0.962$$

Because it is greater than 0.9 80 we reject such messages as spam.

Scanned with CamScanner

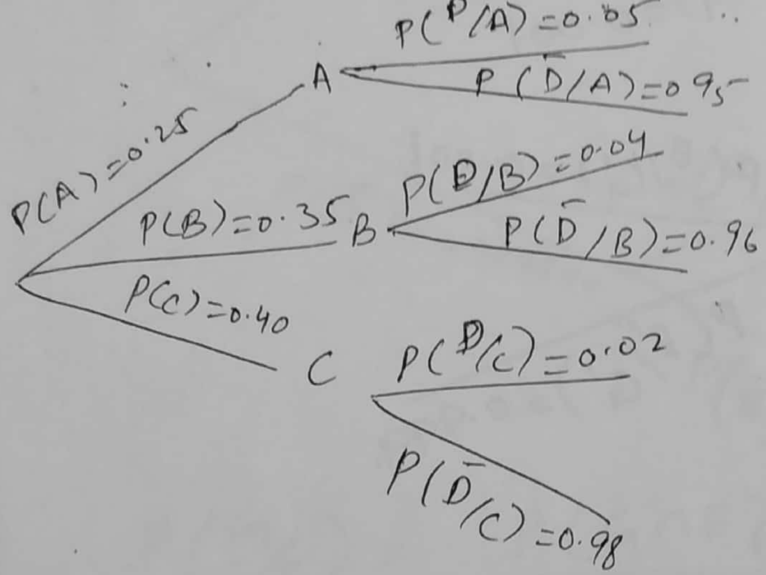
from

Solution:-

$P(A)=0.25,\ P(B)=0.35,\ P(C)=0.40$

Let D: that Bolt is defective

$P(D/A)=0.05 \qquad P(D/B)=0.04 \qquad P(D/C)=0.02$



$$P(C/E) = \frac{P(D/C)\cdot P(C)}{P(A)\cdot P(D/A) + P(B)\cdot P(D/B) + P(C)\cdot P(D/C)}$$

$= 0.232$

Ans