**CS-513 Knowledge Discovery and Data Mining-Midterm-Question 7**

**Poorvi Raut: 20009560**

A COVID test was administered to 1,000,000 individuals. The test correctly identified 95% of those who were sick (P[positive/sick] = 0.95) but also produced a positive result for 10% of those who were not sick (P[positive/not sick] = 0.10). If the prevalence of COVID in this population is 20%, what is the probability that an individual who tested positive is actually sick? What is the probability that an individual who tested negative is actually sick?

Solution:

Let's denote the event that an individual is sick as "S" and the event that an individual tests positive as "P". We are given:

P(S)=0.20 (Covid Prevalence)

P(P|S)=0.95

P(P|~S)=0.10

We have to calculate:

P(S|P) : probability of an individual being sick given that they tested positive

P(S|~P): probability of an individual being sick given that they tested negative

Using Bayes’ Theorem to calculate probability.

P(S|P)=P(P|S)\*P(S)/P(P)

P(S|~P)= P(~ P|S) \* P(S) / P(~ P)

So therefore P(P) = P(P|S) \* P(S) + P(P|~S) \* P(~S)

P(~ P) = P(~P|S) \* P(S) + P(~ P|~ S) \* P(~ S)

Now, we can substitute these values and solve for the probabilities:

P(P) = P(P|S) \* P(S) + P(P|~S) \* P(~S)

= 0.95\*0.20+0.10\*0.80

= 0.215

P(S|P)= P(P|S)\*P(S)/P(P)

= 0.95\*0.20/0.215 = 0.884 approx

P(~ P) = P(~P|S) \* P(S) + P(~ P|~ S) \* P(~ S)

=0.05\*0.20+0.90\*0.80

= 0.785

P(S|~P)= P(~ P|S) \* P(S) / P(~ P)

=0.05\*0.20/0.785

=0.013(approx)

Therefore, the probability that an individual who tested positive is actually sick is approximately 88.4%, and the probability that an individual who tested negative is actually sick is approximately 1.3%.