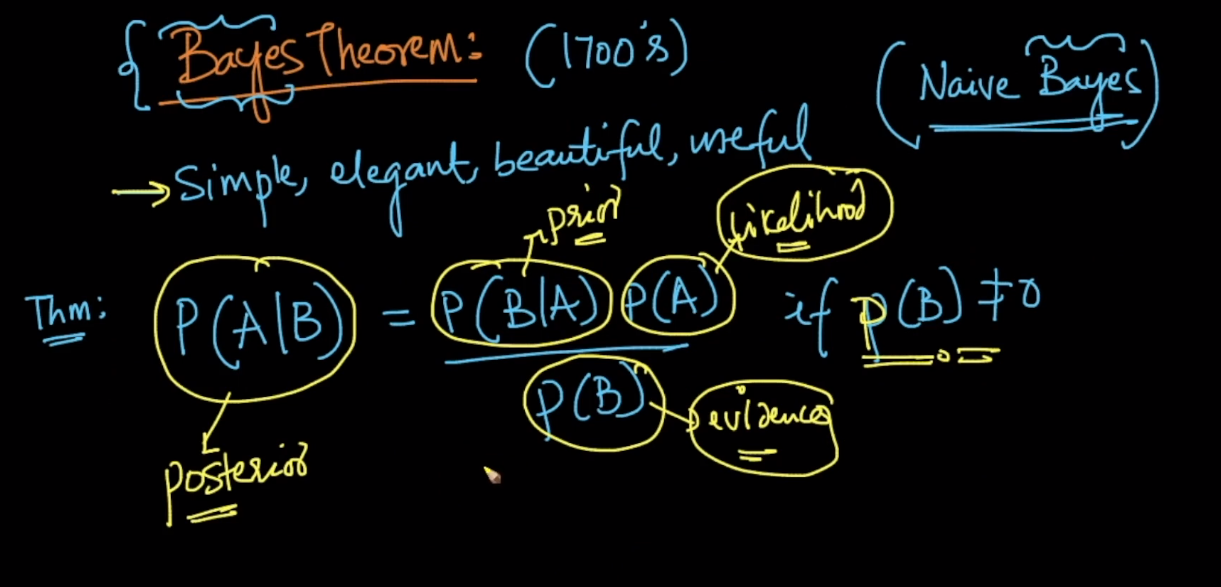
**Bayes Theorem:**

It is one of the most basic , elegant and oldest theorem but very important as Machine Learning point of view because there is an algorithm Naïve Bayes based on the theorem.

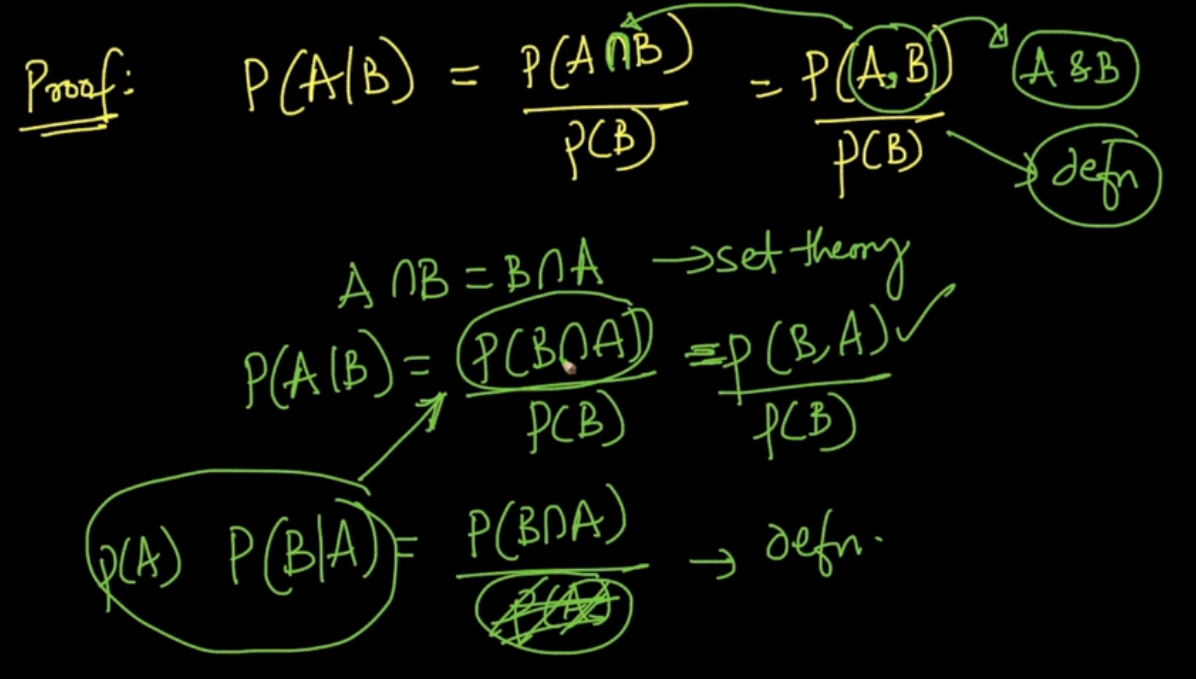
SO the theorem says

P(A|B) = (P(B|A) P(A)) / P(B)

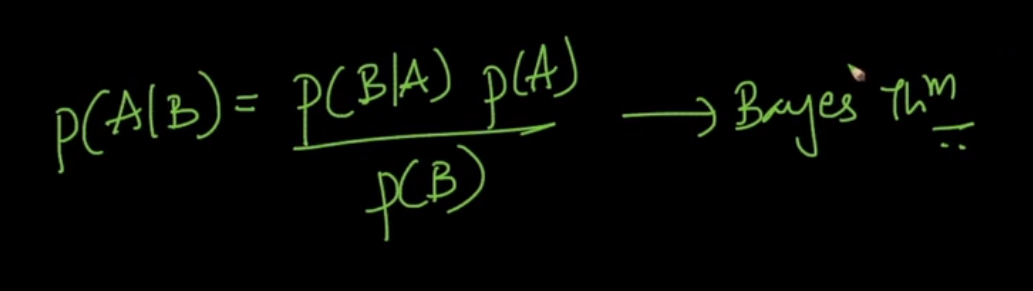
Here each part of equation is given some terminology as shown in image below



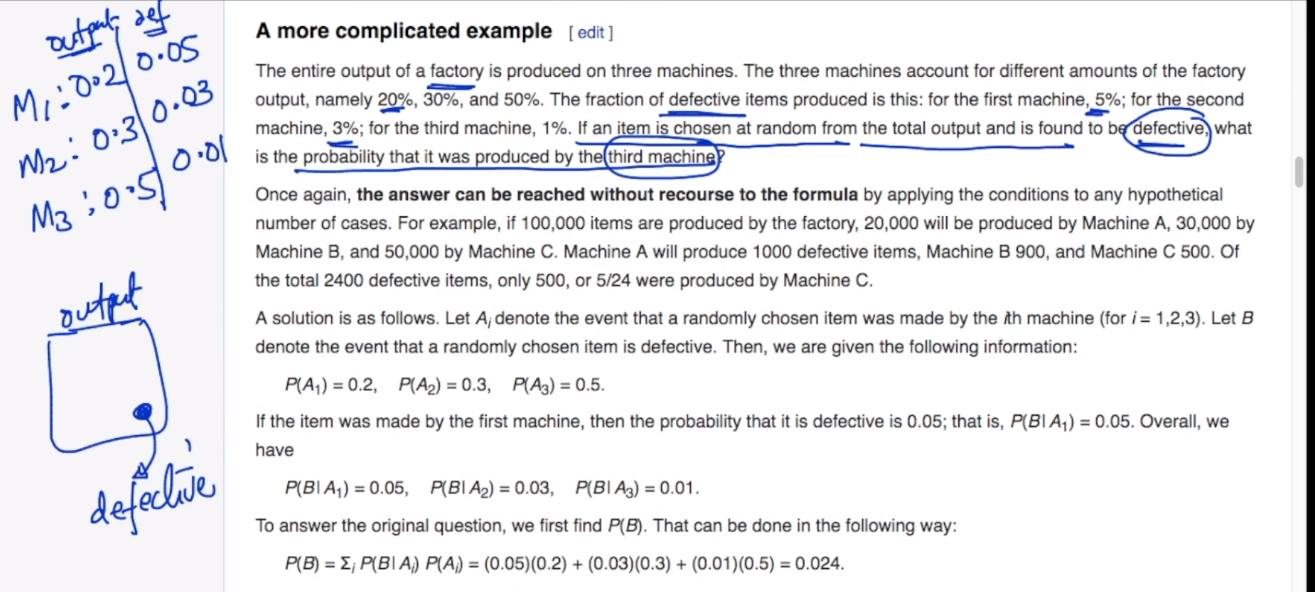
The proof of Bayes Theorem is just a 3 lines proof as shown in below images.



Putting the value of P(B and A) in above equation we get Bayes theorem as shown in below image.



Now lets take a real world example for understanding the theorem more clearly.



Suppose we have 3 machines in a factory and they do work in different proportion i.e., M1 produced 20% , M2 produces 30% and M3 produced 50% of all products

But each machine also produces some defective products M1 = 5% , M2 3% and M3 = 1%

Now the question asked is a product selected at random what is probability that it was produced by M3 .

Since we are already given that one event has happened and asked probability for another event so we know it is conditional probability.

Now we know following things.

P(A1) =P(Product produced by M1) = 0.2

Similarly,

P(A2) = 0.3

P(A3) = 0.5

P(B|A1) = P(Product being defective when it is being produced by M1) = 0.05

Similarly,

P(B|A2) = 0.03

P(B|A3) = 0.01

Now we know that product selected is defective.

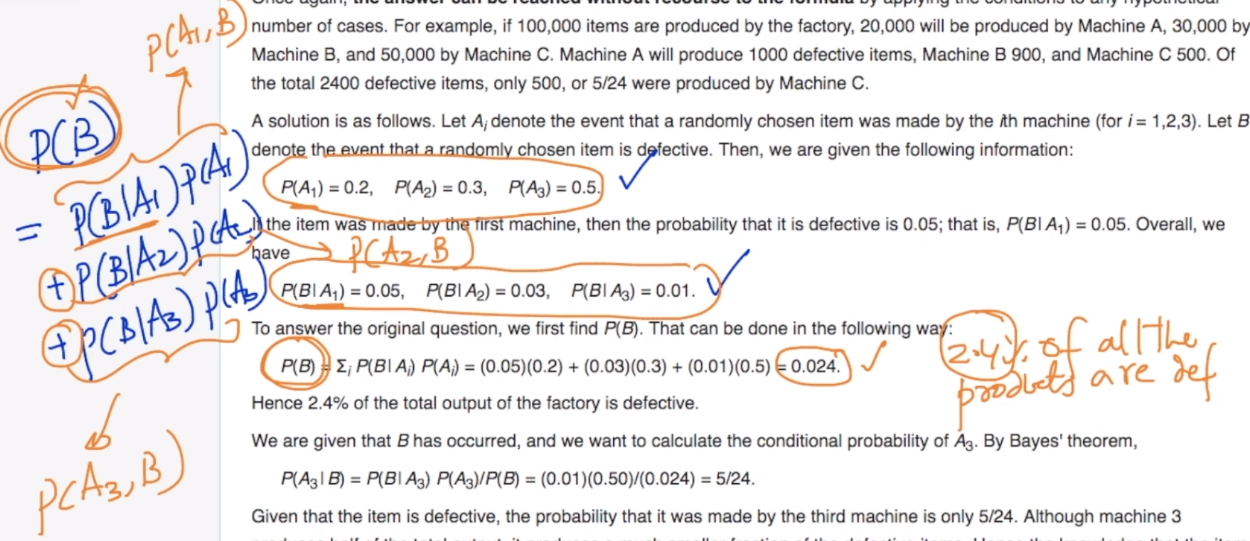
We need to find P(A3) i.e., P(A3|B)

And According to Bayes theorem P(A3|B) = P(B|A3) P(A3) / P(B)

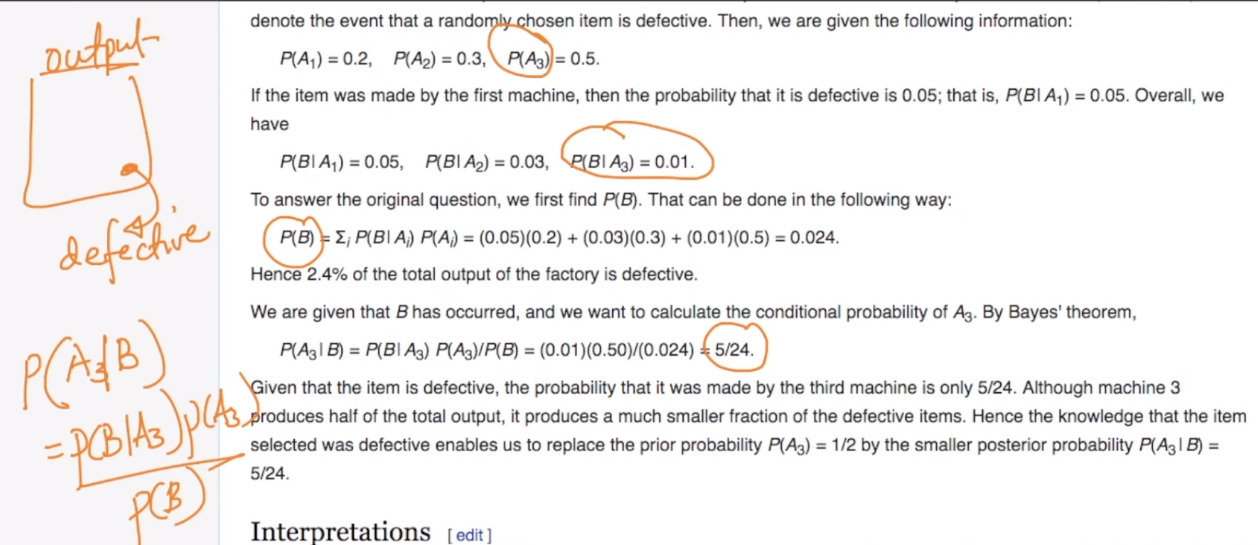
Now we don’t know what is P(B)

SO P(B) = P(A1)\*P(B|A1) + P(A2)\* P(B|A2) +P(A3)\* P(B|A3)

IT means it summation of product being defective when came from machine 1,2,3 i.e., **Summation{i=1 to n} ((P(Ai,B))**



So now we have all the values and by just putting this values in formula of Bayes Theorem we will get our answer for P(A3|B)



Link

<https://www.probabilitycourse.com/chapter1/1_4_2_total_probability.php>