**Naive Bayes**

Before going into the detail of Naive Bayes we have to first understand what is probability

**Probability:** Probability is the measure of the likelihood that an event will occur. Probability is quantified as a number between 0 and 1 (where 0 indicates impossibility and 1 indicates certainty).**[** <https://en.wikipedia.org/wiki/Probability> **]**

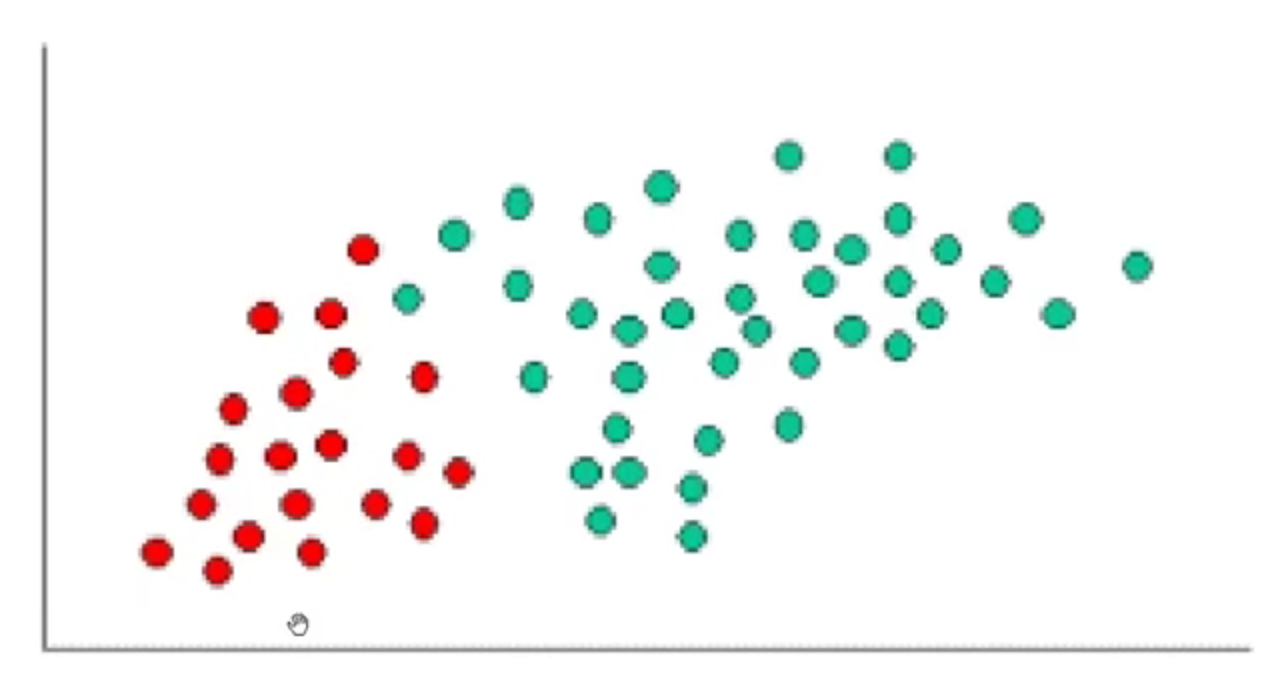
**Prior probability:** A prior probability distribution, often simply called the prior, of an uncertain quantity is the probability distribution that would express one's beliefs about this quantity before some evidence is taken into account.**[** <https://en.wikipedia.org/wiki/Prior_probability> **]**

As for example: One of your friend is talking in a train journey. Now you don't know that to whom he was talking. Now in this scenario if somebody ask you about that whether the person to whom your friend was taking was a man or woman ?

For this case it is a 50/50 probability that the person was a man or a woman. The prior probability for this case ~ their is a 0.5 probability that your friend is talking to a woman and their is a 0.5 probability that your friend is taking to a man.

**Posterior Probability:** Now if you know some information about the conversation, i.e. your friend is talking with a person with long hair, then the probability that the person is a girl will be more. We can say that the probability that the person is a woman is 0.7 and the probability that the person in man will be 0.3. This is called posterior probability.

**Likelihood:**

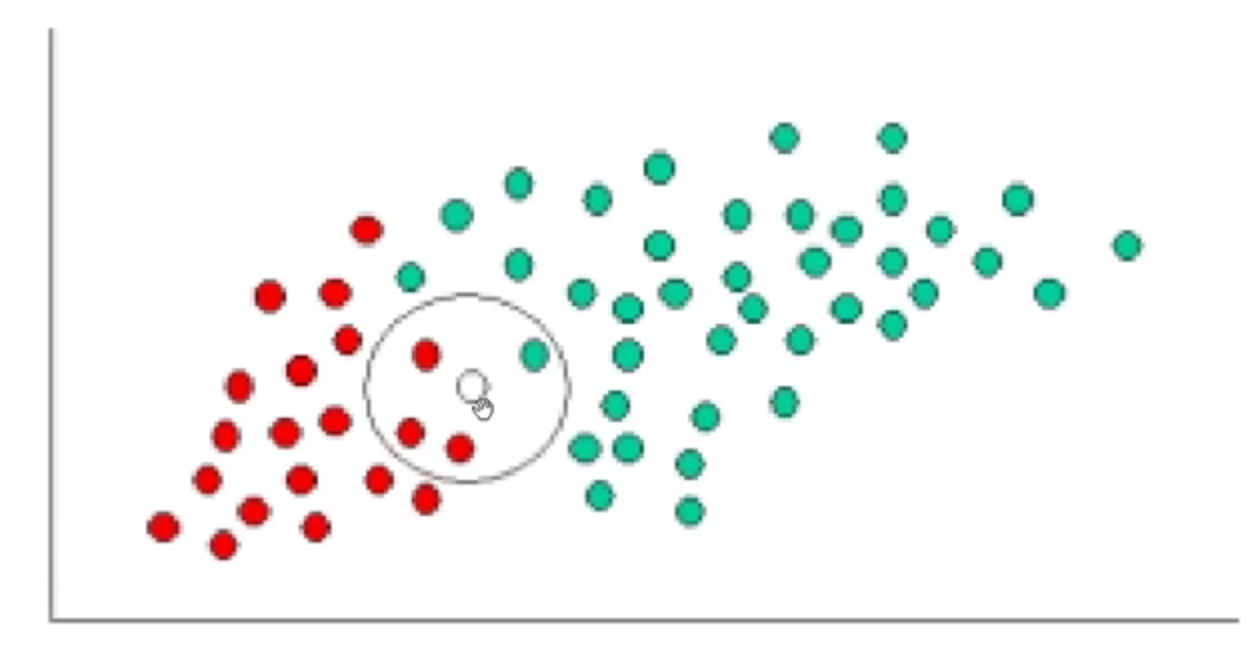


**Fig 1.** Distribution of the two type of data sets. One is red and another is green

**The prior probability of red dots=**

**The prior probability of green dots=**

Now let us consider a new data set come in the vicinity of the distribution as shown in the figure 2. Now we have to make a decision that whether this new point belong to red or green.



**Fig 2.** Presents of new data points

The easiest way to do this is to draw a neighbourhood around the new point. Now the likelihood that the point is red is given by the total number of red point in the neighbourhood divided by the total number of red points in the distribution. Again the likelihood that the point is green given by the total number of green point in the neighbourhood divided by the total number of green points in the distribution.

Likelihood X to be Green =

Likelihood X to be Red =

Say class label **C** has k distinct values

The goal : Given value of all features, we want to predict the probability of C=c1, C=c2….C=ck .

The Bayes Rule: , if