Bayes' Theorem

- Bayes' Theorem is a way of finding a probability when we know certain other probabilities.
- The formula is:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

how often A happens *given that B happens*, written **P(A|B)**,

how often B happens *given that A happens*, written **P(B|A)**

and how likely A is on its own, written **P(A)**

and how likely B is on its own, written **P(B)**

- Let us say P(Fire) means how often there is fire, and P(Smoke) means how often we see smoke, then:

- P(Fire|Smoke) means how often there is fire when we can see smoke

  P(Smoke|Fire) means how often we can see smoke when there is fire

- “Forwards" P(Fire|Smoke) when we know “Backwards" P(Smoke|Fire)

Example:

- dangerous fires are rare (1%)
- but smoke is fairly common (10%) due to barbecues,
- and 90% of dangerous fires make smoke

Solution

We can then discover the probability of dangerous Fire when there is Smoke:

$P(Fire|Smoke) = P(Fire)\ P(Smoke|Fire)/P(Smoke)$

$= 1\% \times 90\%/10\%$

$= 9\%$

So it is still worth checking out any smoke to be sure.

Remembering

- First think "AB AB AB" then remember to group it like: "AB = A BA / B“

$$P(A|B) = \frac{P(A)\,P(B|A)}{P(B)}$$

- The Bayes Rule provides the formula for the probability of Y given X. But, in real-world problems, you typically have multiple X variables.
- When the features are independent, we can extend the Bayes Rule to what is called Naive Bayes.
- It is called ‘Naive’ because of the naive assumption that the X’s are independent of each other. Regardless of its name, it’s a powerful formula.

When there are multiple X variables, we simplify it by *assuming the X's are independent*, so the **Bayes** rule

$$P(Y{=}k \mid X) = \frac{P(X \mid Y{=}k) \, * \, P(Y{=}k)}{P(X)}$$

where, k is a class of Y

becomes, Naive **Bayes**

$$P(Y{=}k \mid X1..Xn) = \frac{P(X1 \mid Y{=}k) * P(X2 \mid Y{=}k) \ldots * P(Xn \mid Y{=}k) \, * \, P(Y{=}k)}{P(X1) * P(X2) \ldots * P(Xn)}$$