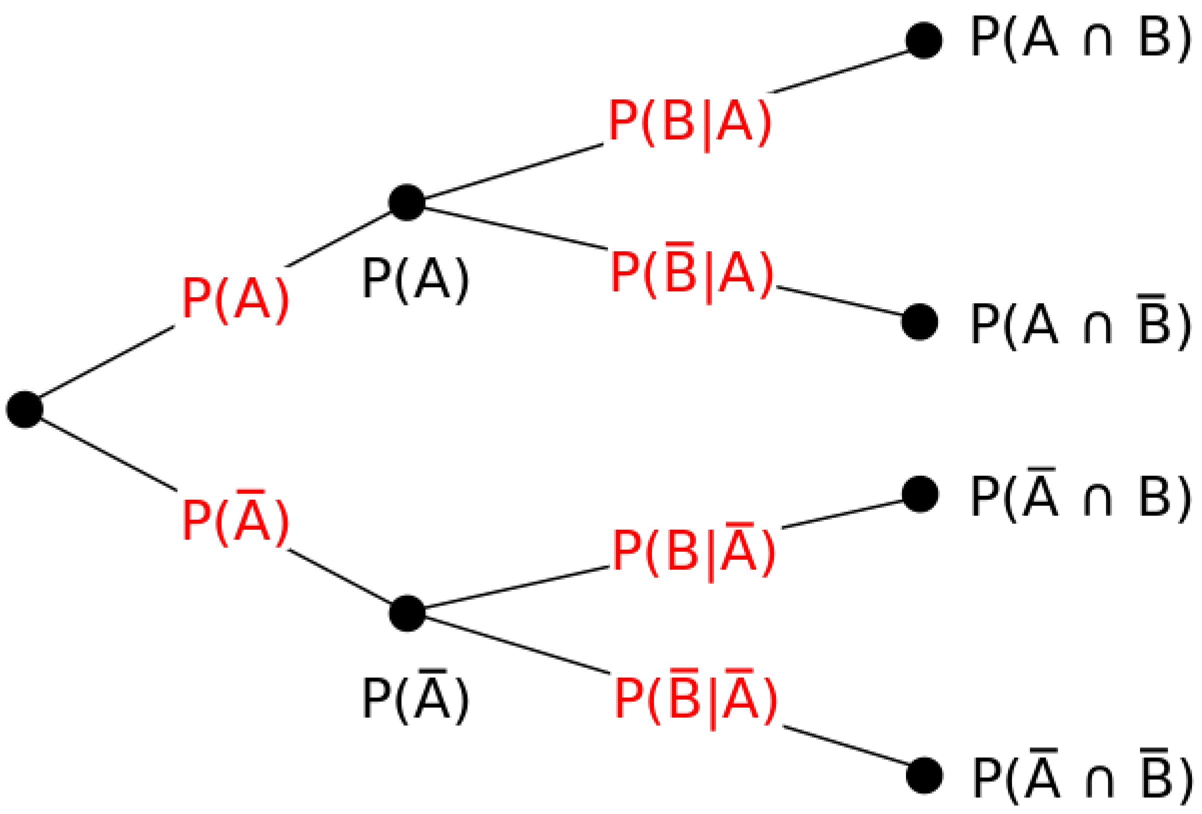
Bayes' Theorem (4A)

Please send corrections (or suggestions) to youngwlim@hotmail.com.

This document was produced by using LibreOffice and Octave.

Intersection Probability



https://en.wikipedia.org/wiki/Conditional_probability

Bayes' Rule (1)

$$P(F|E) = \frac{P(E \cap F)}{P(E)}$$

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

$$P(F|E)P(E) = P(E \cap F)$$

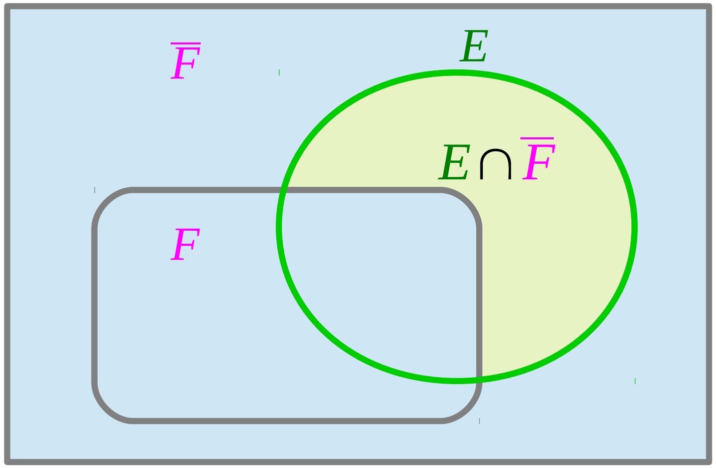
$$P(E|F)P(F) = P(E \cap F)$$

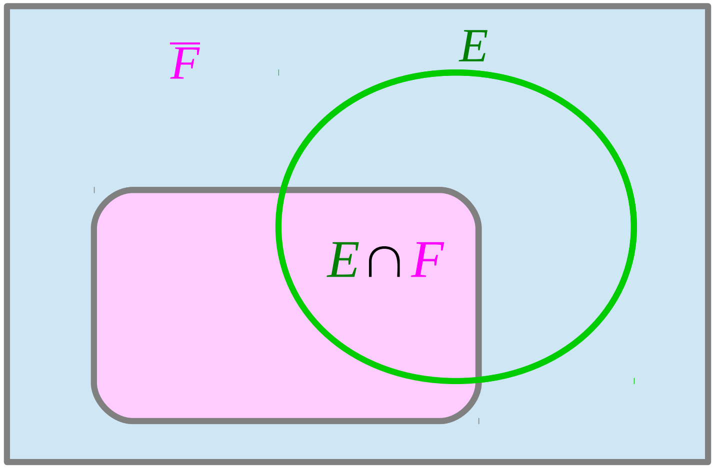
$$P(E \cap F) = P(F|E)P(E) = P(E|F)P(F)$$

$$P(F|E) = \frac{P(E|F)P(F)}{P(E)}$$

$$P(E|F) = \frac{P(F|E)P(E)}{P(F)}$$

$$E = (E \cap \overline{F}) \cup (E \cap F)$$





$$\frac{|E|}{|S|} = \frac{|E \cap \overline{F}|}{|F|} \cdot \frac{|F|}{|S|} + \frac{|E \cap F|}{|F|} \cdot \frac{|F|}{|S|}$$

$$P(E) = P(E|F)P(F) + P(E|\overline{F})P(\overline{F})$$

Bayes' Rule (3)

$$P(F|E) = \frac{P(E|F)P(F)}{P(E)}$$

$$P(E) = P(E|F)P(F) + P(E|\overline{F})P(\overline{F})$$

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|\overline{F})P(\overline{F})}$$

A Priori and a Posteriori

Two types of knowledge, justification, or arguments

A Priori - “from the earlier”

independent of experience

“All bachelors are unmarried”

A Posteriori - “from the later”

Dependent on experience or empirical evidence

“Some bachelors are happy”

Bayes' Rule (1)

| means “given”
H : Hypothesis
E : Evidence

The **prior probability**
the probability of **H**
before **E** is observed.

The **likelihood**
of observing **E** given **H**

$$P(H|E) = \frac{P(E|H) \cdot P(H)}{P(E)}$$

The **posterior probability**
the probability of **H** given **E**,
i.e., after **E** is observed.

the **marginal likelihood**
or "model evidence"
the same for all possible hypotheses

Bayes' Rule (2)

$$P(H|E) = \frac{P(E|H) \cdot P(H)}{P(E)}$$

P(H), the **prior probability** –
the probability of **H** before **E** is observed.
This indicates one's *preconceived beliefs* about how likely different hypotheses are, absent evidence regarding the instance under study.

P(H|E), the **posterior probability** –
the probability of **H** given **E**, i.e., after **E** is observed.
the probability of a hypothesis given the observed evidence

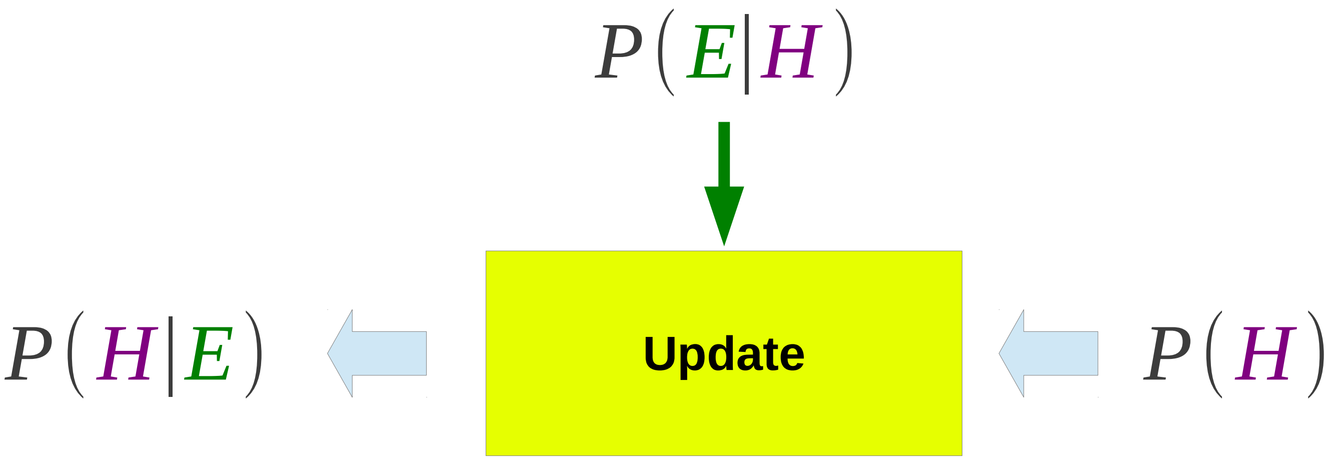
P(E|H), the probability of observing **E** given **H**, is also known as the **likelihood**.
It indicates the compatibility of the evidence with the given hypothesis.

P(E), the **marginal likelihood** or "model evidence". This factor is the same for all possible hypotheses being considered. This means that this factor does not enter into determining the relative probabilities of different hypotheses.

Bayes' Rule (3)

| means “given”
H : Hypothesis
E : Evidence

$$P(H|E) = \frac{P(E|H) \cdot P(H)}{P(E)}$$



The **posterior probability**
the probability of **H** given **E**,
i.e., after **E** is observed.

The **likelihood**
of observing **E** given **H**

The **prior probability**
the probability of **H**
before **E** is observed.

Example

$$P(E|H) \ll P(H) \ll$$

If the **E**vidence doesn't match up with a Hypothesis, one should reject the Hypothesis.
But if a Hypothesis is extremely unlikely a priori, one should also reject it,
even if the **E**vidence does appear to match up.

Three Hypotheses about the nature of a newborn baby of a friend, including:

- H1: the baby is a brown-haired boy
- H2: the baby is a blond-haired girl.
- H3: the baby is a dog.

Consider two scenarios:
I'm presented with **E**vidence in the form of a picture of a blond-haired baby girl.
I find this **E**vidence supports H2 and opposes H1 and H3.

I'm presented with **E**vidence in the form of a picture of a baby dog.
I don't find this **E**vidence supports H3,
since my prior belief in this Hypothesis (that a human can give birth to a dog) is extremely small.

Bayes' rule
a principled way of combining new **E**vidence with prior beliefs, through the application of Bayes' rule.
can be applied iteratively: after observing some **E**vidence, the resulting posterior probability can then be treated as a prior probability, and a new posterior probability computed from new **E**vidence.
Bayesian updating.

Posterior Probability Example (1)

Suppose there are two full bowls of cookies.
Bowl #1 has 10 chocolate chip and 30 plain cookies,
while **bowl #2** has 20 of each.

When **picking a bowl** at random, and then **picking a cookie** at random.
No reason to treat one bowl differently from another, likewise for the cookies.
The drawn cookie turns out to be a plain one.
How probable is it from bowl #1?

more than a half, since there are more plain cookies in bowl #1.

The precise answer

Let H1 correspond to bowl #1, and H2 to bowl #2. P(H1)=P(H2)=0.5.

The event E is the observation of a plain cookie.
From the contents of the bowls, P(E|H1) = 30/40 = 0.75 and P(E|H2) = 20/40 = 0.5.

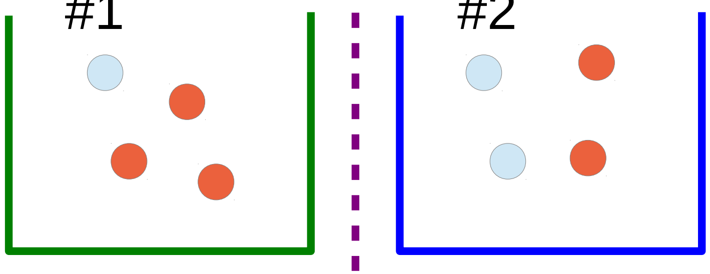
Bayes' formula then yields

$$P(H1|E) = \frac{P(E|H1)\cdot P(H1)}{P(E|H1)P(H1) + P(E|H2)P(H2)} = \frac{0.75 \times 0.5}{0.75 \times 0.5 + 0.5 \times 0.5} = 0.6$$

Posterior Probability Example (2)

Bowl #1

Bowl #2



P(H1)=0.5

P(E|H1) = 30/40 = 0.75

P(H2)=0.5

P(E|H2) = 20/40 = 0.5

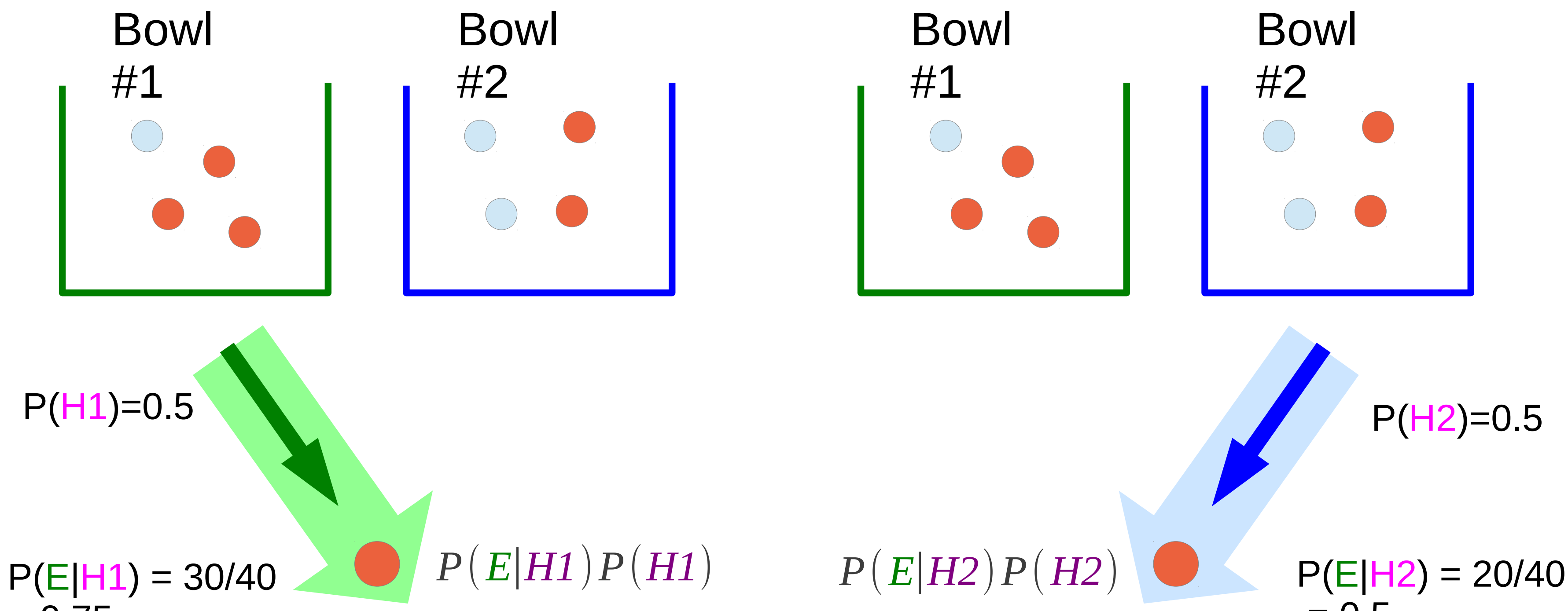
$$P(H1|E) = \frac{P(E|H1) \cdot P(H1)}{P(E|H1)P(H1) + P(E|H2)P(H2)}$$

$$= \frac{0.75 \times 0.5}{0.75 \times 0.5 + 0.5 \times 0.5} = 0.6$$

$$P(E) = P(E|H1)P(H1) + P(E|H2)P(H2)$$

$$P(H1|E) = \frac{P(E|H1) \cdot P(H1)}{P(E)}$$

Posterior Probability Example (3)

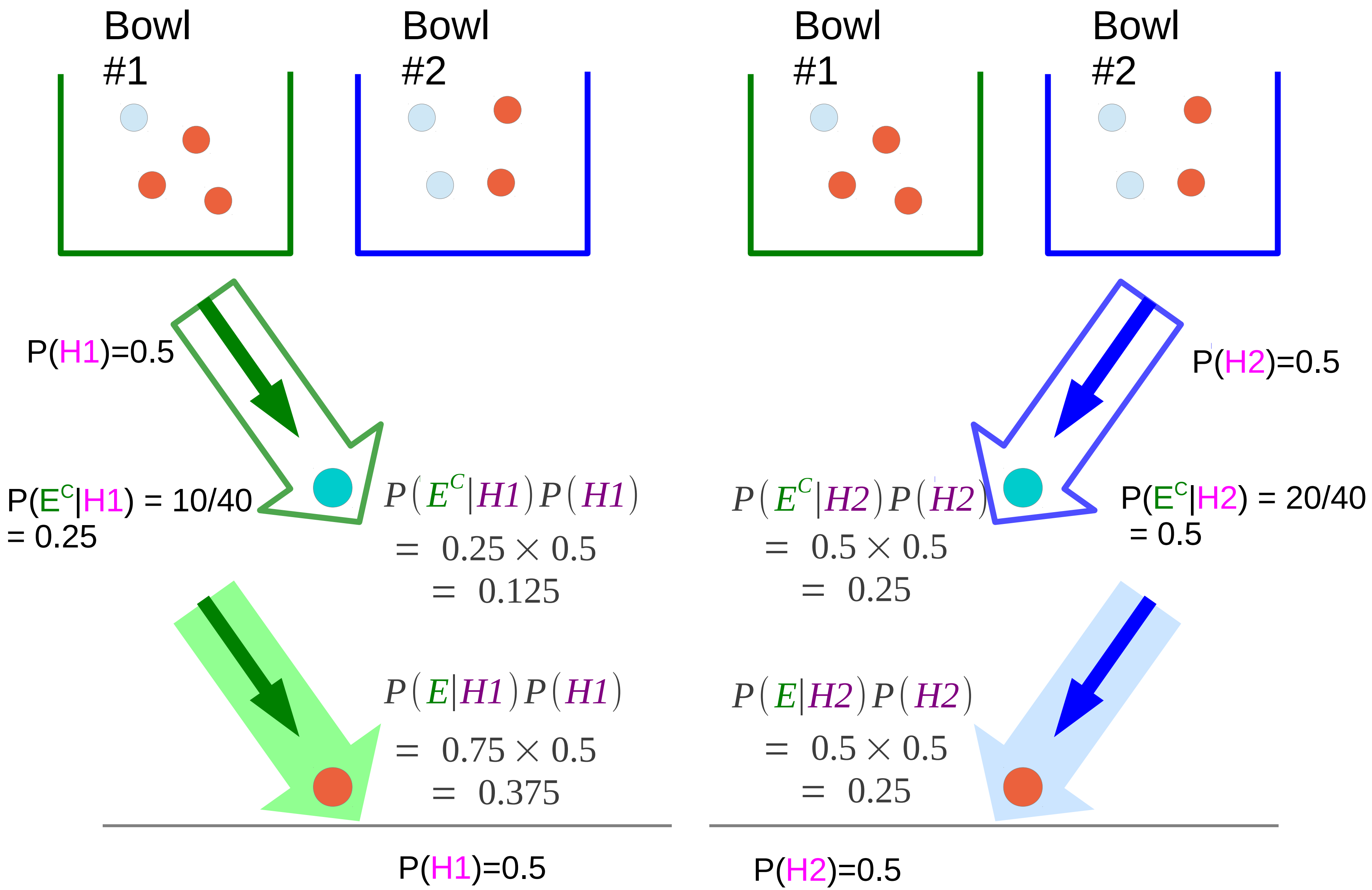


$$P(E|H1)\,P(H1) = 0.75 \times 0.5 = 0.375$$

$$P(E|H2)\,P(H2) = 0.5 \times 0.5 = 0.25$$

$$P(H1|E) = \frac{P(E|H1)\cdot P(H1)}{P(E|H1)\,P(H1) + P(E|H2)\,P(H2)} = \frac{0.75 \times 0.5}{0.75 \times 0.5 + 0.5 \times 0.5} = 0.6$$

Posterior Probability Example (4)



References

[1] http://en.wikipedia.org/
[2] ttps://en.wikiversity.org/wiki/Discrete_Mathematics_in_plain_view
[3] https://en.wikiversity.org/wiki/Understanding_Information_Theory