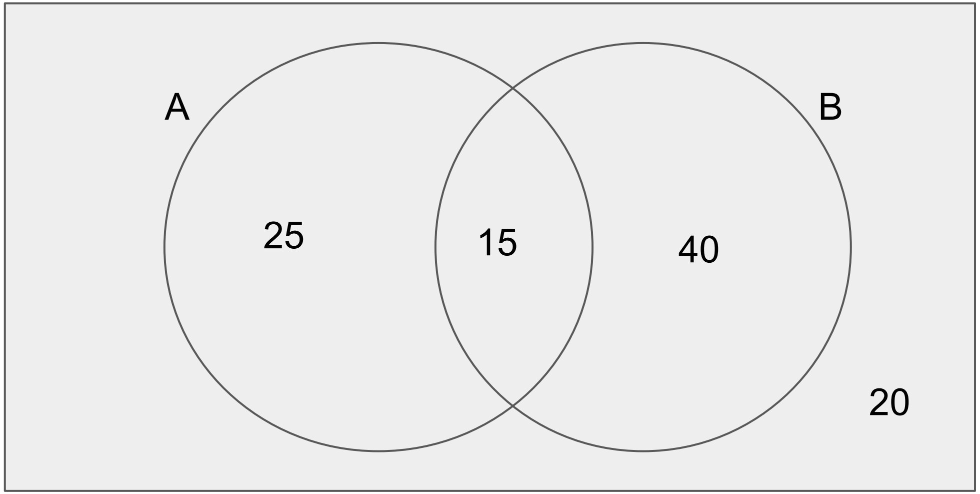
Naive Bayes
Theory and Example

Bayes Theorem



P(A|B) = 15/55 =
(15/100)/(55/100) =
P(A n B)/P(B)

Bayes Theorem

P(A|B) = P(A n B)/P(B) => P(A n B) = P(A|B)xP(B)

P(B|A) = P(A n B)/P(A) => P(A n B) = P(B|A)xP(A)

=> P(A|B)xP(B) = P(B|A)xP(A)   ⇔ P(A|B)/P(A) = P(B|A)/P(B)

Independent Events

A independent of B means:

$P(A|B) = P(A)$ also $P(B|A) = P(B)$

$P(B|A) = P(A \text{ n } B)/P(A)$ => $P(A \text{ n } B) = P(A)xP(B)$

Note also holds when conditioning on a third event C

$P((A \text{ n } B)|C) = P(A|C)xP(B|C)$

Naive Bayes

Want to estimate: P(H=h|E1 = e1,.......,En=en)

Using bayes theorem

P(H=h|E1 = e1,.......,En=en) =
P(E1=e1,.....,En=en|H=h)xP(H=h)/P(E1=e1,.....,En=en)

Difficult to estimate P(E1=e1,.....,En=en|H=h), however if we assume the data is independent (naive) then

P(H=h|E1 = e1,.......,En=en)
≈P(E1=e1|H=h)x,.....,xP(En=en|H=h)xP(H=h)/P(E1=e1,.....,En=en)

Don’t worry about the denominator!!

Naive Bayes Example

H (0 or 1)	E1 (0 or 1)	E2 (0 or 1)
0	0	0
0	0	1
0	1	1
0	0	1
1	0	1
1	1	0

Estimate:

$P(E1=0|H=0) = 3/4$ $P(E1=1|H=0) =1/4$ $P(E1=0|H=1) = 1/2$ $P(E1=1|H=1) = 1/2$
$P(E2=0|H=0) = 1/4$ $P(E2=1|H=0) = 3/4$ $P(E2=0|H=1) = 1/2$ $P(E2=1|H=1) = 1/2$
$P(H = 0) = 2/3$ and $P(H = 1) = 1/3$

Naive Bayes Example

Given E1 = 0 and E2 = 1 what is more probably H = 0 or H = 1?

Which is higher $P(H=1|E1 = 0,E2 = 1)$ or $P(H=0|E1 = 0,E2 = 1)$?

Applying estimate derived previously

$P(H=1|E1 = 0,E2 = 1) = P(H = 1)xP(E1 = 0|H = 1)xP(E2 = 1|H = 1)/P(E1 = 0,E2 = 1)$

Similarly

$P(H=0|E1 = 0,E2 = 1) = P(H = 0)xP(E1 = 0|H = 0)xP(E2 = 1|H = 0)/P(E1 = 0,E2 = 1)$

Plugging in the Numbers Gives

$P(H=1|E1 = 0,E2 = 1) = (1/3x1/2 \text{ x } 1/2)/P((E1 = 0,E2 = 1)) = (1/12)/P(E1 = 0,E2 = 1))$

$P(H=0|E1 = 0,E2 = 1)) = (2/3x3/4 \text{ x}3/4)/P((E1 = 0,E2 = 1)) = (3/8)/P((E1 = 0,E2 = 1))$

As denominator the same for both H=0 is the most likely given the evidence