Bayes Theorem

$$P(A/B) = \frac{P(B/A) \cdot P(A)}{P(B)}$$

B is evidence

A is Hypothesis

Assumption is that features are independent

$\Rightarrow$ that's why it is naive.

$$P(y|X) = \frac{P(X|y)\, P(y)}{P(X)}$$

$$X = (x_1, x_2, \ldots, x_n)$$

$\Rightarrow$

$$P(y|x_1 \ldots x_n) = \frac{P(x_1|y)\, P(x_2|y) \ldots P(x_n|y) \cdot P(y)}{P(x_1)\, P(x_2) \ldots P(x_n)}$$

Now Dr is static, so

$$P(y|x_1, x_2, \ldots x_n) \propto P(y) \prod_{i=1}^{n} P(x_i|y)$$

We need to find class y with max. prob.

$$y = \underset{y}{\operatorname{argmax}}\ P(y) \prod_{i=1}^{n} P(x_i|y)$$

Now

$$\underbrace{P(y|X)}_{\text{Posterior}} = \frac{\overbrace{P(X|y)}^{\text{Likelihood}} \cdot \overbrace{P(y)}^{\text{Prior Prob class}}}{\underbrace{P(X)}_{\text{Prior Prob features}}}$$

<u>Ex</u> Table → Weather → Play

Sunny – No, Overcast → Yes, Rainy – Yes,
Sunny – Yes, Sunny – Yes, Overcast – Yes,
Rainy – No, Rainy – No, Sunny – Yes,
Rainy – Yes, Sunny – No, Overcast – Yes

Overcast - yes, Rainy - no

Problem

Players will play if weather is sunny

$$P(Yes/Sunny) = \frac{P(Sunny/Yes) * P(Yes)}{P(Sunny)}$$

$$= \frac{3/9 * 9/14}{5/14}$$

$$= 0.60 \Rightarrow$$ Ans ✓

$\Rightarrow$

$$P(no/Sunny) = 0.40$$

Another Example

Persons $\Rightarrow$ A and B — Both an office same time

Info $P(A) = 0.5$ $P(B) = 0.5$

The person we saw is with red sweater

A wears red 2 times/week

B " " 3 " "

Now

Initial Prob is $P(A) = P(B) = 0.5$

then we have new information

We get Posterior Prob

$P(A) = 2/5 = 0.4$

$P(B) = 3/5 = 0.6$

Prior + Information $\Rightarrow$ Posterior

Known Prob   Inferred Prob

$P(\text{A wears Red})$ $P\left(\text{Person with red is A}\right)$

$P(\text{B wears Red})$ $P\left(\text{person with red is B}\right)$

Known   Inferred

$P(A)$

$P(R|A)$ $\xrightarrow[\text{Thm.}]{\text{Bayes}}$ $P(A|R)$

R → Related Event

A → Event

Again same example

Now lets say we know

A comes 3 times a week and
B comes 1 time a week

Now Prior Probabilities are

$P(A) = 3/4 = 0.75$
$P(B) = 1/4 = 0.25$

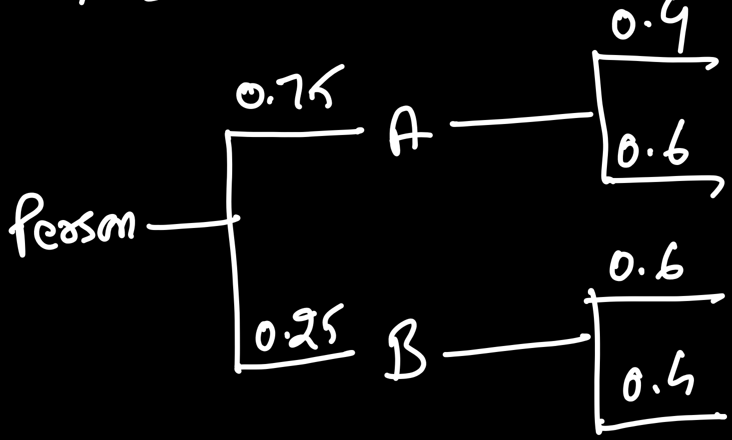
Knowledge $\Rightarrow$ Person had a <u>Red</u> sweater

<u>Rule</u>
0.4 $\leftarrow$ A wears red 2 times
0.6 $\leftarrow$ B wears red 3 times

Naively, we think probs are not
0.75 / 0.25 but should be close to
each other, since B wears more.

Now we want to compute

Posterior Probs



$R \quad 0.75 \cdot 0.4 = 0.3$

$R' \quad 0.75 \cdot 0.6 = 0.45$

$R \quad 0.25 \cdot 0.6 = 0.15$

$R' \quad 0.25 \cdot 0.4 = 0.1$

$\underline{\overline{\Sigma = 1}}$

Person we saw was wearing Red

so our sample space is reduced

<u>Universe</u>

0.3 , 0.15

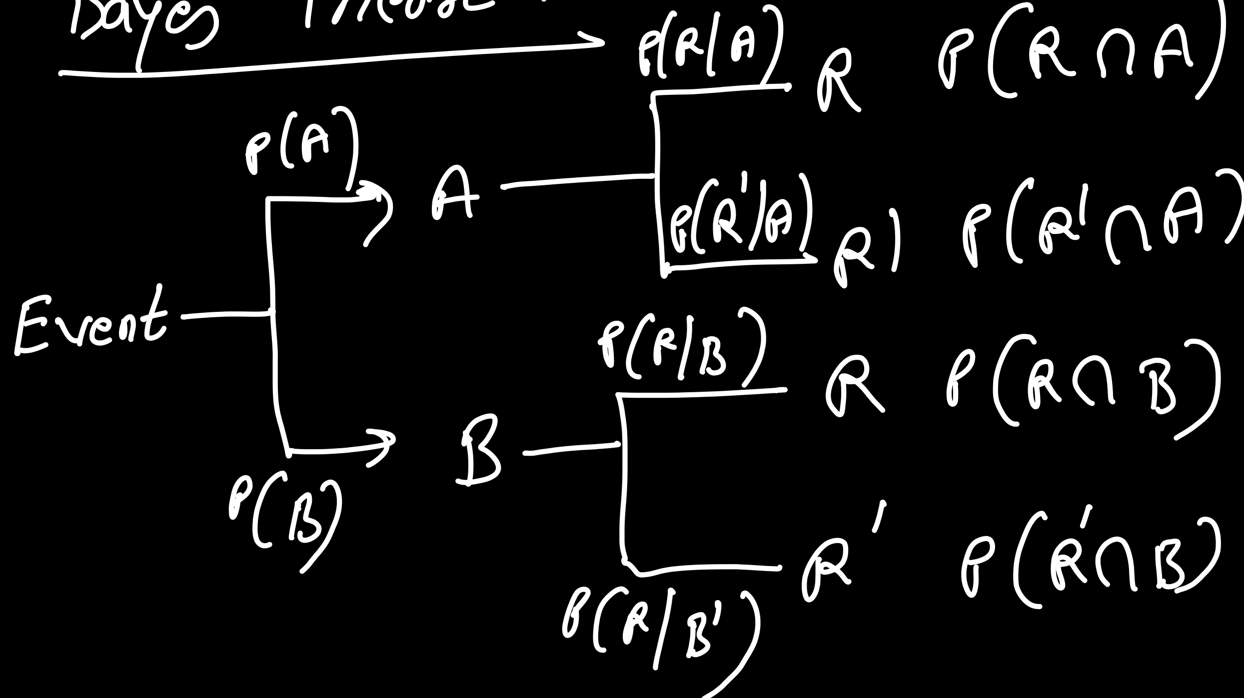
We have prob and normalize

$$P(A/R) = \frac{0.3}{0.3+0.15} = 0.67$$

Bayes

$$P(B|R) = \frac{0.15}{0.3+0.15} = 0.33$$ Thm

Bayes Theorem



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

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

$P(A) . P(R/A) + P(B) . P(R/B)$

Prior Probabilities are

$P(A)$ and $P(B)$

and after event R

Posterior Probs are

$P(A/R)$ and $P(B/R)$

Applications

Doctor recommends a test to a patient that has 99 % accuracy

$\Rightarrow$ 99 % Sensitivity Detects 99 sick out of 100 sick

99 % Specificity Detects 99 healthy of 100 health

Data say 1 out of 1000 suffer from disease

Now Test is +ve

Now Prob of being sick when test is +ve

<u>Bayes Theorem</u>

S → Sick   H → Healthy

+ → positive

$P(S) = \frac{1}{1000} = 0.0001$

$P(H) = 0.9999$

$P(+|S) = 0.99$

$$P(+|H) = 0.01$$

$$P(S|+) = \frac{P(S)\,P(+|S)}{P(S)\,P(+|S) + P(H)\,P(+|H)}$$

$$= \frac{0.0001 * 0.99}{0.0001 * 0.99 + 0.9999 \times 0.01}$$

$$= \frac{0.000099}{0.000099 + 0.009999}$$

$$= \frac{0.000099}{0.010098} = 0.0098039 2$$

$< 1\%$

Now explore

1M
1,000,000

- S — 100
  - + — 99
  - − — 1
- H — 999,900
  - + — 9999
  - − — 989901

1 out of 10,000 → S

Now if Test is positive

$$S \cap + = 99$$

$$H \cap + = 9999$$

$$P(\text{sick} \mid +\text{ve}) = \frac{99}{99 + 9999} \approx 0.0098$$

$< 1\%$

healthy group of healthy people +ve so much larger ⇒

False Positive

Test fails 1% time but this is much much larger

1 out of 10,000 sick

$$\Rightarrow \frac{1}{10,000} \times 100 = \frac{1}{100} = 0.01\% \text{ Sick}$$

1% error $\Rightarrow$

100 out of 10,000 be diagnoised sick

but only 1 will be sick from 10,000

Naive Bayes and ML

Apply Naive Bayes repeatedly

Ex Spam vs Not-Spam

Spam	Ham/Not-Spam
E1 → Win Money Now	E1 → How are you
E2 → Make Cash Easy	E2 → There you are

E3 → Cheap Money Reply | E3 → Can I borrow money

E4 → Say Hi to Grandma

E5 → Was the exam easy

New Email → Easy Money

Now

$$p(A|R) = \frac{p(R|A) \cdot P(A)}{p(R|A) \cdot p(A) + p(R|A^c)\, p(A^c)}$$

$B = A^c$

Now if A and R are independent

⇒ Naive Assumption

$p(R/A) = p(R)$

$Dr = p(R)\, p(A) + p(R)\, p(A^c)$

$= p(R)\left[\underbrace{A(A) + p(A^c)}_{1}\right]$

$= p(R)$

Thus

$$p(A/R) = \frac{p(R/A) \cdot p(A)}{p(R)}$$

$$p(A/R) \propto p(R/A) \cdot p(A)$$

$$p(span/EM) \propto p(EM/span) * p(span)$$

Naive Assum $p(A \cap B) = p(A)\, p(B)$

$$p(span/EM) \propto p(E/span) \cdot p(M/span) \cdot p(span)$$

$$\propto \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{3}{8}$$

lly

$$p(ham/EM) \propto p(E/ham)\, p(M/ham)\, p(ham)$$

$$\propto \frac{1}{5} \cdot \frac{1}{5} \cdot \frac{5}{8}$$

Thus

$$p(\text{spam}/Em) \propto \frac{1}{12} = p_1$$

$$p(\text{ham}/Em) \propto \frac{1}{40} = p_2$$

Since the email must be either spam or ham so these should add to 1

So we normaliz

$$\text{Total} = \frac{1}{12} + \frac{1}{40} = \frac{40+12}{40.12} = \frac{\overset{13}{\cancel{\overset{26}{\cancel{52}}}}}{40.\underset{3}{\cancel{\underset{6}{\cancel{12}}}}}$$

$$= 13/120 = s$$

$$p(\text{spam}/Em) = \frac{1/12}{13/120} = \frac{1}{12} \propto \frac{120}{13}$$

$$= 10/13 \checkmark$$

$$\text{||ly } p(\text{ham}/Em) = 3/13 \checkmark$$

Extending Naive Bayes

$$p(spam|(easy, money, cheap)) = p(S|EMC)$$

$$\propto p(EMC|spam)\, p(spam)$$

Thus

$$p(S|EMC) \propto p(E|S)\, p(m|S)\, p(C|S)\, p(S)$$

|||ly

$$p(H|EMC) \propto p(E|H)\, p(m|H)\, p(C|H)\, p(H)$$

Then normalize