Tutorial 03

1 Exercise.

a) $P(\text{Tennis} = \text{Yes}) = \underbrace{\dfrac{9}{14}}_{0.64}$ → We only need to count the proportion of "Yes" over the total number of instances

b) $\underbrace{P(\text{Outlook} = \text{rain} / \text{Tennis} = \text{Yes})}_{\text{likelihood}}$ → we only need to count the proportion of instances wher tennis is Yes. Then among these instances count the ones where outlook is rain

$$= \frac{3}{9} = 0.33$$

2 Exercise.

So what this is asking is

$$P(\text{Tennis} = \text{Yes} \mid \text{Rain}, \text{Hot}, \text{Normal}, \text{Weak})$$

↓

We cannot get this directly. Let's try to use the bayes rule to simplify it.

Bayes Rule: $P(A|B) = \dfrac{P(A) \cdot P(B|A)}{P(B)}$

OR

$$\frac{P(A \cap B)}{P(B)}$$

So

$$P(Yes|Rain, hot, Normal, Weak) = \frac{\overbrace{P(Rain, hot, Normal, Weak|Yes)}^{Likelihood} \times \overbrace{P(Yes)}^{Yes}}{P(rain, hot, Normal, Weak)}$$

The likelihood is almost impossible, small, or hard to obtain with just counting. To solve this we use Naive Bayes. This allows to get an approximate likelihood that asumes that the variables are independent.

$P(Rain, hot, Normal, Weak|Yes) =$

$$\underbrace{P(Rain|Yes) \times P(hot|Yes) \times P(Normal|Yes) \times P(Weak|Yes)}_{\text{We can get this probabilities with just counting the proportion of the attribute with the instances with Yes.}}$$

$3/9 \times 2/9 \times 6/9 \times 6/9 = 0.0329$

So now, we can calculate the probability of Yes.

$P(Yes) = 9/14$

Then we have that

$$P(Yes|Rain, hot, Normal, Weak) = \frac{9/14 \times 0.0329}{P(Rain, hot, Normal, Weak)}$$

however the probability distribution has to ~~add~~ sum up to one. So ~~we~~ then we can see what is the most probable outcome "Yes" or "No". So we have to do the same thing but for "No".

So

~~P~~ also written as $\neg$ tennis

$P(No \mid rain, hot, normal, weak) =$ ~~No~~ Bayes Rule

$$\frac{P(rain, hot, normal, weak \mid No) \times P(No)}{P(rain, hot, normal, weak)}$$

$P(No) = 5/14$

$P(rain, hot, normal, weak / No) =$

$P(rain / No) \times P(hot / No) \times P(normal / No) \times P(weak / No)$

$2/5 \times 2/5 \times 1/5 \times 2/5 = 0.0128$

So we have

is lower than the counterpart

$P(No / rain, hot, normal, weak) = 0.0128 \times 5/14 = \boxed{0.00457}$

And we know that $P(No / rain, hot, normal, weak) +$ $P(Yes / rain, normal, weak) = 1$. So we can calculate the ~~actual~~ normalized probabilities

$P(Yes / rain, hot, normal, weak) = \dfrac{0.021}{0.021 + 0.0045} = \boxed{0.821}$

normalized

$P(Yes / \ldots)$ $P(No / \ldots)$

So this is more probabl

Also another way to se this is

Not normalized probabilities

$$P(Yes \mid rain, hot, normal, weak) = \frac{0.021}{\cancel{P(rain, hot, normal, weak)}}$$

delete mutual

$$P(No \mid rain, hot, normal, weak) = \frac{0.00457}{\cancel{P(rain, hot, normal, weak)}}$$

Still $P(Yes \mid rain, hot, normal, weak)$, is bigger without normalisation.

Question 4:

Is exactly the same Question as previous one....
but asked in a different way.

answer

$V_{NB} = \underset{Tennis = Yes, No}{\arg\max}\ P(tennis) \cdot P(rain, hot, normal, weak / tennis)$

So we have to check wether the probability of Yes is higher than No or viceversa. We did this before

for Yes =

$P(Yes) \cdot P(rain/Yes) \cdot P(hot/Yes) \cdot P(normal/Yes) \cdot P(weak/Yes) =$

$\boxed{0.021}$ → higher

$P(No) \cdot P(rain/No) \cdot P(hot/No) \cdot P(normal/No) \cdot P(weak/No) =$

0.00457

Tutorial 3

Question 4.

Is almost the same as previous exercises.... But there is a special thing you should take care. This exercise deals with the ocassion when one of the components of the likelihood is 0.

$V_{NB} = \underset{\text{Tennis} = \text{Yes, No}}{\arg\max} \; P(\text{tennis}) \cdot P(\text{cloud, hot, normal, weak} \mid \text{Tennis})$

So, we have to do the calculations for Yes and No, and see which one is higher.

$P(\text{Yes}) \cdot P(\text{cloud, hot, normal, weak} \mid \text{Yes})$

$9/14 \cdot P(\text{cloud} \mid \text{Yes}) \cdot P(\text{hot} \mid \text{Yes}) \cdot P(\text{normal} \mid \text{Yes}) \cdot P(\text{weak} \mid \text{Yes})$

$\underbrace{9/14 \cdot 4/9 \cdot 4/9 \cdot 6/9 \cdot 6/9}_{0.282}$

This is a problem...
We need to use th m-estimate to solve it.

m-estimate $\Rightarrow P(A/B) = \dfrac{n_c + m \times p}{n + m}$ $P(\ldots$

$P(\text{No}) \cdot P(\text{cloud, hot, normal, weak} \mid \text{No})$

$5/14 \cdot P(\text{cloud} \mid \text{No}) \cdot P(\text{hot} \mid \text{No}) \cdot P(\text{normal} \mid \text{No}) \cdot P(\text{weak} \mid \text{No})$

$\to 0$

there are no instances with this case

Tutorial 3

So we need to apply th m-estimate to p(cloud | No).

$$p(\text{cloud} \mid \text{No}) = \frac{\overset{n_c}{0} + \overset{p}{\frac{1}{3}} \times \overset{m}{5}}{\underset{m}{5} + \underset{n}{5}} = 0.166$$

So now we can compare both approaches

$$\underbrace{\underset{5/14}{P(\text{No})} \cdot \underset{0.166}{P(\text{cloud} \mid \text{No})} \cdot \underset{2/5}{P(\text{hot} \mid \text{No})} \cdot \underset{1/5}{P(\text{normal} \mid \text{No})} \cdot \underset{2/5}{P(\text{weak} \mid \text{No})}}_{0.00189}$$

P(Yes | cloud, hot, normal, weak) is higher

Tutorial 3

Question 5.

$p(x) = \underbrace{0.4}\ \underbrace{N_1(x)} + 0.6\, N_2(x)$

Gaussian 1

mean = 3

variance = 5

Gaussian 2

mean = 4

variance = 2

We just need to use the formula with the weights.....

x = 6

$$\underbrace{\underset{\uparrow \atop 6}{p(x)} = 0.4\ \frac{-1}{e^{2\times 5}}(6-3)^2 + 0.6\ \frac{-1}{e^{2\times 2}}(6-4)^2}_{0.383}$$

Tutorial @3
Question 6.

Solution of the operations are on Keats .....