Module 2

Probabilistic learning

Events are possible outcomes such as a head or tail result in a coin flip.

A trial is a single oppurtunity for the event to occur such as a coin flip.

Probability

- The probability of an event can be estimated by

$$= \frac{\text{no. of trials in which an event occured}}{\text{total no. of trials}}$$

- Notation P(A) is used to denote probability.
- Total probability of all possible outcomes of a trial must always be 100%.
- Chances of occuring an event → Probability
- Event with defined outcomes → Random Experiment
- Set of all possible outcome can be listed → Sample Space
- $P = 1 - P(P')$
- Mutually Exclusive → When two Events are mutually exclusive, it means that they cannot both occur at the same time.
- If two Events are Exhaustive, it means that one of them must occur.
- If two Events are totally unrelated, they are called independent Events.

3 unbaised coins are tossed
a) write sample space
b) find i) P(all heads) ii) P(atleast 2 heads) iii) P(atmost 2 h[illegible]

- a) { (HHH) (HTH) (HHT) (HTT) (THH) (TTH) (THT) (TTT) }

b) i) $1/8$

ii) $4/8 = 1/2$

iii) 0 to 2 heads $= 7/8$

Q. A bag contains 7 red & 4 white balls. 2 balls are drawn at random. Find
i) P(both balls are red)
ii) P(one is white, other is red)
iii)

- i) $\frac{7C2}{11C2}$ // ii) $\frac{4C_1 \times 7C_1}{11C_2}$

Q. 3 bulbs are chosen at random from 15 bulbs of which 5 are defective. Find
i) P(none is defective)
ii) P(exactly one is defective)
iii) P(atleast one is defective)

total = 15
def = 5
not = 10

- i) $\frac{10C3}{15C3}$, ii) $\frac{5C_1 \times 10C_2}{15C3}$ iii)

iii) atleast one is defective = 1 def 2 non
2 def 1 non
3 def.

$$= \frac{5C_1 \times 10C_2}{15C_3} + \frac{5C_2 \times 10C_1}{15C_3} + \frac{5C_3}{15C_3}$$

Q. A drawyer contains 50 bolts & 150 nuts in a mechanical workshop. Half of bolts & half of the nuts are rusted. If one of them is chosen at random what is the Prob(it is rusted item or a bolt)

50 bolts
150 nuts
Rusted :
25 bolts
75 nuts

- Two Events : drawing a rusted item
drawing a bolt item.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Rusted item = 25 + 75 = 100 //

Total item = 50 + 150 = 200 //

P(drawing rusted item) = 100/200 = $\underline{1/2}$

P(drawing bolt) = $\frac{50}{200} = 1/4$ //

P(rusted & bolt) = $P(A \cap B) = 25/200$

P(rusted or bolt) = $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$$= \frac{1}{2} + \frac{1}{4} - \frac{25}{200} = \frac{6}{8} - \frac{25}{200}$$

$$= \underline{5/8}$$

Q. The lifespan of 5000 electrical components are measured to access their reliability. The lifespan (L) is recorded & results are shown below. Find the prob. that a randomly selected component will last

i) more than 4 year

ii) b/w 4 & 6 year

iii) less than 5 year

Life Span	No.
A) $L > 6$	600
B) $5 < L \leq 6$	2250
C) $4 < L \leq 5$	1850
D) $L \leq 4$	300

i) more than 4 year = $P(A) + P(B) + P(C)$

$$= \frac{600 + 2250 + 1850}{5000} = \frac{4700}{5000} = \frac{47}{50}$$

ii) ~~more~~ b/w 4 & 6 = $P(B) + P(C)$

$$= \frac{2250 + 1850}{5000} = \frac{4100}{5000} = \frac{41}{50}$$

iii) less than 5 year - $P(C) + P(D)$

$$= \frac{1850 + 300}{5000} = \frac{2150}{5000} = \frac{43}{100}$$

Conditional probability with Baye's theorem

* The relationship b/w dependent events can be described using Baye's theorem.
* The notation $P(A|B)$ can be read as the probability of event A given that event B occured. This is known as conditional Probability. Since the probability of A is dependent on what happened with event B.

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

Note : $\boxed{P(A/B) = \frac{P(A \cap B)}{P(B)} \qquad P(B/A) = \frac{P(A \cap B)}{P(A)}}$

$$\boxed{\begin{aligned} P(A \cap B) &= P(A/B)P(B) \\ &= P(B/A)P(A) \end{aligned}}$$

Proof : $P(A) = \frac{n(A)}{S}$, $P(B) = \frac{n(B)}{S}$, $P(A \cap B) = \frac{n(A \cap B)}{S}$

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

$$= \frac{n(A \cap B)}{S} \cdot \frac{S}{n(B)}$$

$$\boxed{P(A|B) = P(A \cap B) \cdot P(B)}$$

III^{ly} $P(B|A) = P(A \cap B) \cdot P(A)$

Q. A bag contains 3 red & 4 white balls. Two draws are made without replacement. What is the P(both are red).

$P(A) = P($first ball is red$)$

$P(B|A) = P($Second ball is red without replacement$)$

$$P(A \cap B) = P(A) \cdot P(B/A)$$

$$= \cancel{3C} \frac{3}{7} \cdot \frac{2}{6}$$

$$= 1/7$$

8. 3 baskets are given each containing R & W balls as given below

B1	B2	B3
6R	2R	1R
4W	6W	8W

A basket is selected at random & a ball is drawn from it. The ball is red. Find the prob. that the selected basket is B1.

e_1 = select basket
e_2 = select ball

- E_1 : Choosing basket 1

E_2 : Choosing basket 2

E_3 : Choosing basket 3

$P(E_1) = P(E_2) = P(E_3) = 1/3$

A be the event of drawing a red ball

red B1

$P(A|E_1) = 6/10 = 3/5$

$P(A|E_2) = 2/8 = 1/4$

$P(A|E_3) = 1/9$

~~B1~~ $|A$

$E_1|A$

Basket

$$P(E_1|A) = \frac{P(E_1 \cap A)}{P(E_1 \cap A) + P(E_2 \cap A) + P(E_3 \cap A)}$$

Basket 1

$$\boxed{= \frac{P(E_1 \cap A)}{\sum_{i=1}^{3} P(A|E_i)\,P(E_i)} = \frac{P(E_1) \cdot P(A|E_1)}{\sum_{i=1}^{3} P(A|E_i)\,P(E_i)}}$$

$$= \frac{1/3 \cdot 3/5}{1/3 \cdot 3/5 + 1/3 \cdot 1/4 + 1/3 \cdot 1/9}$$

$$= \frac{1/5}{1/5 + 1/12 + 1/27} = \frac{1/5}{519/1620} = \frac{324}{519}$$

Q. A basket is chosen at random & 2 balls are drawn without replacement from the same basket. If both balls are red then what is the P(ball was selected from B_1 under this condition). What ~~P(prob~~ P(B3 is selected).

- A : be the both ball are red

$P(A|E_1) = \frac{6C2}{10C2}$ $P(A/E_2) = \frac{2C2}{8C2}$ $P(A/E_3) = 0$

$$P(E_1/A) = \frac{\cancel{P(E_1)\ P(E_1 \cap A)}\ P(E_1) \cdot P(A|E_1)}{P(E_1) \cdot P(A/E_1) + P(E_2) P(A|E_2) + P(E_3) P(A/E_3)}$$

$$= \frac{1/3 \cdot \frac{6C2}{10C2}}{\frac{1}{3} \cdot \frac{6C2}{10C2} + \frac{1}{3} \cdot \frac{2C2}{8C2} + 0}$$

$$= \frac{1/9}{1/9 + 1/28} = \frac{1/9}{37/228} = \frac{228}{333}$$

7/12/2021

Bayesian Theorem on Probability

If $E_1, E_2 \dots E_n$ are mutually exclusive events with $P(E_i) \neq 0$ for $i = 1, 2, \dots n$. Then for any Event 'A', that the subset of $\bigcup_{i=1}^{n} E_i$ (all the events) such that $P(A) > 0$ ($P(A) \neq 0$), then we have $\boxed{P(E_i|A) = \frac{P(E_i) \cdot P(A|E_i)}{\sum_{i=1}^{n} P(E_i) \cdot P(A|E_i)}}$

Bag contains 5 balls, 2 balls are drawn & are found to be white. What is the P(all the balls are white) ?

$P(AW \mid \underset{\text{2 white}}{\widetilde{2w}})$

all white

$$P(E \mid A) = \frac{P(AW) \cdot P(2w \mid AW)}{\text{total}}$$

white

5, 4, 3, 2 — (4)

E_1 is the event representing bag with 2 white

$E_2 \rightarrow$ 3 white

$E_3 \rightarrow$ 4 white

$E_4 \rightarrow$ 5 white

$P(E_1) = P(E_2) = P(E_3) = P(E_4) = \frac{1}{4}$ (equally likely) / total

$P(E_1) = P(A \mid E_1) = \frac{2C_2}{5C_2}$ (selecting two balls) $\frac{2\times1}{5\times4/2} = 1/10$

$P(A \mid E_2) = \frac{3C_2}{5C_2} = 3/10$ $\quad \frac{3\times2}{5\times4/2}$

$P(A \mid E_3) = \frac{4C_2}{5C_2} = 2/5$ $\quad \frac{4\times2}{5\times4}$

$P(A \mid E_4) = \frac{5C_2}{5C_2} = 1$

Q. A Bag contains 7 red, & 3 black marbles & another bag contains 4 red & 5 black marbles. One marble is transferred by first bag to second bag & then a marble is taken out of second bag at random. If this marble happens to be red. Find the P(black marble was transferred).

$$P(B \mid R) = \frac{P(B) \cdot P(R/B)}{P(B) \cdot P(R/B) + P(R) \cdot P(R/R)}$$

Red

black being transferd

7R + 3B → 1

4R + 5B → 2

$P(B) = 3/10 \qquad P(R) = 7/10$

$P(R/B) = 4/10$ (after adding Bk to B_2) 9+1 = 10

$P(R/R) = 5/10$ (after adding red to B_2) 4+1

$$P(B/R) = \frac{\frac{3}{10} \times \frac{4}{10}}{\left(\frac{3}{10} \times \frac{4}{10}\right) + \left(\frac{7}{10} \times \frac{5}{10}\right)} = \frac{12}{12 + 35} = \frac{12}{47}$$

Q. The members of consulting firm hire cars or vehicles from rental agencies x, y & z as 60%, 30% & 10% respectively. If 9%, 20% & 6% of the vehicles from agencies x, y & z turn up. & if a rental car delivered to the firm doesnot turn up. what is the P(it came from agency Y).

$$P(y/\bar{T}) = \frac{P(y) \cdot P(\bar{T}/y)}{P(x) \cdot P(\bar{T}/x) + P(y) \cdot P(\bar{T}/y) + P(z) \cdot P(\bar{T}/z)}$$

delivered by y

doesn't Turn up.

$P(y) = \frac{30}{100} = \frac{3}{10}$

$P(x) = \frac{60}{100} = 6/10$

$P(z) = \frac{10}{100} = 1/10$

$\cancel{P(x/T)} =$

$P(\bar{T}/x) = 91\% = \frac{91}{100}$

$P(\bar{T}/y) = 80\% = 80/100$

$P(\bar{T}/z) = 94\% = 94/100$

$$P(y/\bar{T}) = \frac{3/10 \times 80/100}{\frac{6}{10}\times\frac{91}{100} + \frac{3}{10}\times\frac{80}{100} + \frac{1}{10}\times\frac{94}{100}}$$

$$= \frac{3\times 80}{6\times 91 + 3\times 80 + 1\times 94} =$$

Q. In a ~~family~~ factory, machine x & y produce springs of the same type: of this production, machine x & y produce 5% & 10% defective spring respectively. Machine x & y produce 40% & 60% of total O/P of the factory. One spring is selected at random & is found to be defective. What is the P(defective spring was produced by machine x) ?

~~P(x)~~

$P(\cancel{x}_D) = \frac{5}{100}$ ~~$P(y_D) = \frac{10}{100}$ (defective~~) $P(x) = \frac{40}{100}$ (x)

$P(y) = \frac{60}{100}$ (y)

$P(x_D/D) = \cancel{\dfrac{P(x_D) \cdot P(D/x_D)}{P(x_D)P(D/x_D) + P(y_D) \cdot P(D/y_D)}}$ $P(D/x) = \frac{5}{100}$ (x is def)

$P(D/y) = \frac{10}{100}$ (y is def)

$P(\cancel{D/x_D}) =$

$$P(x|D) = \frac{P(x)\cdot P(D/x)}{P(x)\cdot P(D/x) + P(y)\cdot P(D/y)}$$

$$= \frac{\frac{40}{100}\times\frac{5}{100}}{\frac{40\times 5}{100\times 100} + \frac{60}{100}\times\frac{10}{100}}$$

$= 1/2$ //

~~$\frac{40\times 5}{100\times 100}$~~

~~$\frac{40}{100}$ 800~~

~~100×100~~

200

600

Q. Assume that the word 'OFFER' occurs in 80% of the spam messages. Also let us assume offer occur in 10% of the desired " . If 30% of the received emails are considered as spam, & if I will receive a new message which contain the word 'OFFER'. What is the P(it is a spam)?

$$\cancel{p(s)} = P(O/S)_{\text{spam}}^{\text{offer}} = \frac{80}{100}$$

$$\underset{\text{offer desired}}{P(O/D)} = \frac{10}{100}$$

$$P(S) = \frac{30}{100}$$

$$P(D) = \frac{70}{100}$$

$$P(S/O) = \frac{P(S) \times P(O/S)}{P(S) \times P(O/S) + P(D) \times P(O/D)}$$

$$= \frac{\frac{30}{100} \times \frac{80}{100}}{\frac{30}{100} \times \frac{80}{100} + \frac{70}{100} \times \frac{10}{100}}$$

$$= \frac{30 \times 80}{100 \times 100} \times \frac{100 \times 100}{2400}$$

$$= \frac{24}{31}$$

$$P(S/O) = \frac{P(S) \times P(O/S)}{P(S) \times P(O/S) + P(D) \times P(O/D)}$$

$$= \frac{\frac{30}{100} \times \frac{80}{100}}{\frac{30 \times 80}{100 \times 100} + \frac{70}{100} \times \frac{10}{100}}$$

$$\frac{\frac{2400}{10000}}{\frac{2400}{10000} + \frac{700}{10000}}$$

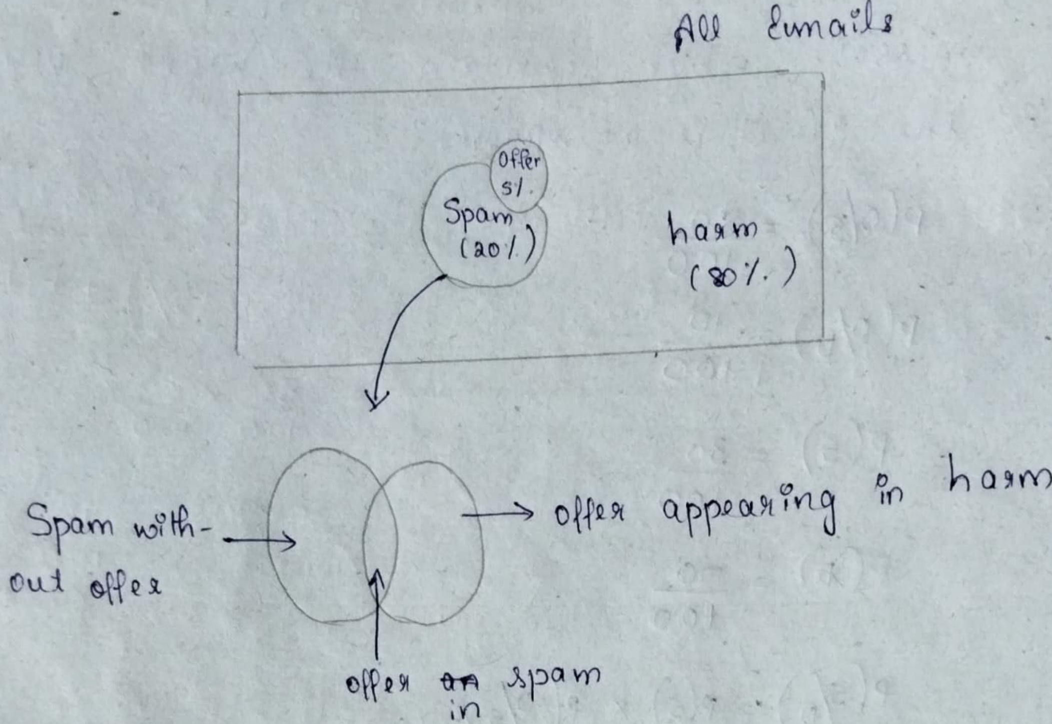
$$= \cancel{\frac{2400}{10000} \times \frac{10000}{3100}}$$

JOINT PROBABILITY

Consider the following Venn diagram:

'offer' circle doesn't completely fill the spam circle, nor is it completely contained by the spam circle.

This implies that not all the spam messages contain the word 'offer' & not every email ~~...~~ with the word 'offer' is spam.



- Calculating P(spam ∩ offer) depends on Joint Probability of the two events or how the probability of one event is related to the probability of the other.
- To understand how Baye's theorem works in pra suppose that you were tasked with guessing the prob. that an incoming email was spam.
- Without any additional evidence, the most reasonab guess would be the prob. that any prior message was spam (20%). This estimate is known as the

PRIOR PROBABILITY

- Suppose you were told that the incoming message used the term 'offer'. The prob. that the word 'offer' was used in previous spam message is called the LIKELIHOOD & the prob. that offer appeared in any message at all is known as the MARGINAL LIKELIHOOD.

- By applying Baye's theorem to this evidence, we can compute a POSTERIOR PROBABILITY that measures how likely the message is to be spam.
- If the posterior prob. is ~~>60%~~ >50%, the msg is more likely to be spam than harm & it should be filtered.

likelihood

$$P(spam/offer) = \frac{P(offer|spam)\ P(spam)}{P(offer)}$$

Posterior probability

Prior probability

Marginal likelihood

- To calculate the components of Baye's theorem, we must construct a frequency table that records the no. of times 'offer' appeared in spam & harm message.
- The frequency table can then be used to construct a likelihood table

	Offer		
Freq.	Yes	No	Total
Spam	4	16	20
Harm	1	79	80
Total	5	95	100

	Offer		
Likelihood	Yes	No	Total
Spam	4/20	16/20	20
Harm	1/80	79/80	80
Total	5/100	95/100	100

THE NAIVE BAYES ALGORITHM

* The naive Bayes (NB) algorithm describes a simple app^n using Baye's theorem for classification.

> The naive Bayes algorithm is named as such because it makes a couple of 'naive' assumptions about the data.

> In particular, naive Bayes assumes that all of the features in the dataset are equally important & independent. Those assumptions are rarely true in most of the real-word applications.

> Eg: If you were attempting to identify Spam by monitoring email msg, it is almost certainly true that some features will be more important than others.

> Naive Bayes algorithm is a supervised learning algo., which is based on Bayes theorem & used for solving classification problems.

The Naive Bayes classification

Eg: The naive Bayes learner is trained by constructing a likelihood table for the appearance of 4 words (w_1, w_2, w_3 & w_4) for 100 emails.

	OFFER(w_1)		MONEY(w_2)		LOAN(w_3)		UNSUBSCRIBE (w_4)	
LIKELIHOOD	YES	NO	YES	NO	YES	NO	YES	NO
Spam	4/20	16/20	10/20	10/20	0/20	20/20	12/20	8/20
Ham	1/80	79/80	14/80	66/80	8/80	72/80	23/80	51/80
Total	5/100	95/100	24/100	76/100	8/80	92/100	35/100	65/100

calculate the prob that the message is a spam given that offer = YES, MONEY = NO, loan = NO, unsubsribed = YES.

An

$$P\left(S / w_1 \cap \neg w_2 \cap \neg w_3 \cap w_4\right) = \frac{P(S) \cdot P(w_1 \cap \neg w_2 \cap \neg w_3 \cap w_4 | S)}{P(S) \cdot P(w_1 \cap \neg w_2 \cap \neg w_3 \cap w_4 | S) + P(h) \cdot P(w_1 \cap \neg w_2 \cap \neg w_3 \cap w_4 | h)}$$

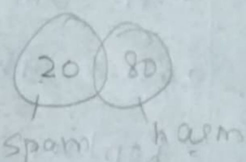
$$= \frac{T_1}{T_1 + T_2}$$

$$T_1 = P(S) \cdot P(w_1/S) \; P(\neg w_2/S) \; P(\neg w_3/S) \; P(w_4/S)$$

$$= \frac{20}{100} \times \frac{4}{20} \times \frac{10}{20} \times \frac{20}{20} \times \frac{12}{20}$$

P(S) = prior prob

$P(S) = \frac{20}{100}$



$$= \frac{6}{20 \times 25} = \frac{3}{250}$$

$$T_2 = P(H) \cdot P(w_1/H) \; P(\neg w_2/H) \; P(\neg w_3/H) \; P(w_4/H)$$

$$= \frac{80}{100} \times \frac{1}{80} \times \frac{66}{80} \times \frac{72}{80} \times \frac{23}{80}$$

$$= \frac{13662}{6400000} = 0.002$$

$$\text{Desired Probability} = \frac{0.012}{0.002 + 0.012} = 0.857$$

THE LAPLACE ESTIMATOR

Consider the above problem, if we have to find the $P(S | w_1 \cap w_2 \cap w_3 \cap w_4)$

Clearly likelihood of spam is,

$$(4/20) \times (10/20) \times (0/20) \times (12/20) \times (20/100) = 0$$

$$\therefore \; P(\text{spam}) \text{ is } \frac{0}{0 + 0.009} = 0$$

$$P(\text{harm}) \text{ is } \frac{0.00005}{0 + 0.00005} = \underline{\underline{1}}$$

- This results suggest that the msg is spam with 0% & harm with 100% probability. It is therefore very likely that the msg has been incorrectly classified.
- This problem might arise if an event never occur for one or more levels of the class.
- A soln. to this problem involves using something called the Laplace Estimator.
- Laplace estimator essentially adds a small no. to each of the counts in the frequency table, which ensures that each feature has a non-zero probability of occurancy with each class.
- Typically, laplace estimator is set to 1, which ensures that each class-feature combination is found in the data atleast once.

20/12/2021

Q. Given the following data, on a set of patients seen by a doctor. Can the doctor conclude that a person having chills, fever, mild headache & without running nose has flu.

Chills	Running nose	Headache	Fever	has flu?
Y	N	Mild	Y	N
Y	Y	No	N	Y
Y	N	Strong	Y	Y
N	Y	Mild	Y	Y
N	N	No	N	N
N	Y	Strong	Y	Y
N	Y	Strong	N	N
Y	Y	Mild	Y	Y

- Here, it is given categorical ~~data~~ features & features with categorical values.

(instance given in Question)
↳ chills, fever, mild headache & without running nose.

$P\left(\text{Flu} / \text{chills}=Y, RN=N, HA=\text{mild}, \text{Fever}=Y\right) =$ (using Bayesian theorem)

$$= \frac{P(\text{Flu}) \cdot P(\text{chills}|\text{flu}) \cdot (RN=N/\text{flu}) \cdot P(HA=\text{mild}|\text{flu}) \cdot P(\text{fever}|\text{flu})}{\text{Total Probability}}$$

$$= \frac{P(\text{Flu}) \cdot P(\text{chills}/\text{flu}) P(RN=N|\text{flu}) \cdot P(HA=\text{mild}|\text{flu}) \cdot P(\text{Fever}|\text{Flu})}{\text{Total Prob} + P(\text{Flu}=N)}$$

$P(\text{Flu}) = 5/8$ $\quad$ $P(\text{Flu}=N) = 3/8$

$P(chills/Flu) = \frac{3 \text{ (chills)}}{5 \text{ (Flu = Yes)}}$ $P(chills = N | Flu = 0) = 1/3$

$P(RN = N | Flu) = \frac{1}{5}$ $P(RN = N | Flu = 0) = 2/3$

$P(HA = mild | Flu) = 2/5$ $P(HA = mild | Flu = 0) = 1/3$

$P(fever = Y | Flu) = 4/5$ $P(fever = Y | Flu = 0) = 1/3$

$$Pf = \frac{\frac{5}{8} \times \frac{3}{5} \times \frac{1}{5} \times \frac{2}{5} \times \frac{4}{5}}{\left(\frac{5}{8} \times \frac{3}{5} \times \frac{1}{5} \times \frac{2}{5} \times \frac{4}{5}\right) + \left(\frac{3}{8} \times \frac{1}{3} \times \frac{2}{3} \times \frac{1}{3} \times \frac{1}{3}\right)}$$

$$= \frac{\frac{3}{125}}{\frac{3}{125} + \frac{1}{108}} = \frac{3}{125} \times \frac{125 \times 108}{(3 \times 108) + 125}$$

$$= \frac{3 \times 108}{(3 \times 108) + 125} = \frac{324}{449}$$

Q.

Example	color	Type	Origin	Stolen
1	Red	Sports	Domestic	Y
2	Red	Sports	Domestic	N
3	Red	Sports	Domestic	Y
4	~~Red~~ Yell	Sports	Domestic	N
5	Yellow	Sports	Imported	Y
6	Yellow	SUV	Imported	N
7	Yellow	SUV	Imported	Y
8	Yellow	SUV	Domestic	N
9	Red	SUV	Imported	N
10	Red	Sports	Imported	Y

Using Bayesian algorithm, determine whether a red domestic SUV car is stolen or not using the Above data.

$$P(\text{stolen} \mid \text{Color}=\text{Red}, \text{Type}=\text{SUV}, \text{Origin}=\text{Domestic}) =$$

$$\frac{P(\text{stolen}) \cdot P(\text{Red}/\text{stolen}) \cdot P(\text{SUV}/\text{stolen}) \; P(\text{Domestic}/\text{stolen})}{\text{Case 1} + P(\text{stolen}=N) \cdot P(\text{Red}/\text{stolen}=N) \cdot P(\text{SUV}/\text{stolen}=N) \; P(\text{Dom}/\text{stolen}=N)}$$

$P(\text{stolen}) = 5/10$ $\qquad$ $P(\text{not stolen}) = 5/10$

$P(\text{Red}/\text{stolen}) = 3/5$ $\qquad$ $P(\text{Red}/\text{stolen}=N) = 2/5$

$P(\text{SUV}/\text{stolen}) = 1/5$ $\qquad$ $P(\text{SUV}/\text{stolen}=N) = 3/5$

$P(\text{Dom}/\text{stolen}) = 2/5$ $\qquad$ $P(\text{Dom}/\text{stolen}=N) = 3/5$

$$= \frac{\cancel{\frac{5}{10}} \times \cancel{\frac{3}{5}} \times \frac{1}{\cancel{5}} \times \frac{2}{\cancel{5}}}{\left(\cancel{\frac{5}{10}} \times \cancel{\frac{3}{5}} \times \frac{1}{\cancel{5}} \times \frac{2}{\cancel{5}}\right) + \left(\cancel{\frac{5}{10}} \times \frac{2}{\cancel{5}} \times \frac{3}{\cancel{5}} \times \frac{3}{5}\right)} = \frac{1}{1+3} = \frac{1}{4}$$

$$= \frac{\cancel{2 \times 3 \times 1 \times 2}}{5 \times \cancel{5 \times 5} \times 5} \times \frac{\cancel{(5 \times 5 \times 5 \times 5)(5 \times 5 \times 5 \times 5)}}{}$$

NUMERICAL FEATURES IN BAYES

Eg: consider the spam email problem, bucket ← Bin Binning

Sir, dear ⟶ This is not considered as features. some significant words such as unsubscribed, offer.. etc are called features.

Categorical features (P(offer included or not ?)

when time is