PROBLEM #1

(a) $P(\text{fruit}) = \frac{2}{5}$

$P(\neg\,\text{fruit}) = \frac{3}{5}$

$P(\text{orange} \mid \text{fruit}) = \frac{1+1}{6+9} = \frac{2}{15}$

$P(\text{orange} \mid \neg\,\text{fruit}) = \frac{4+1}{11+9} = \frac{5}{20}$

$P(\text{Mango} \mid \text{fruit}) = \frac{1+1}{6+9} = \frac{2}{15}$

$P(\text{Mango} \mid \neg\,\text{fruit}) = \frac{0+1}{11+9} = \frac{1}{20}$

$P(\text{Melon} \mid \text{fruit}) = \frac{0+1}{7+10} = \frac{1}{17}$

$P(\text{Melon} \mid \neg\,\text{fruit}) = \frac{0+1}{12+10} = \frac{1}{22}$

$$P(\text{Red} \mid \text{fruit}) = \frac{0+1}{6+9} = \frac{1}{15}$$

$$P(\text{Red} \mid \neg\text{fruit}) = \frac{2+1}{11+9} = \frac{3}{20}$$

$$P(\text{lemon} \mid \text{fruit}) = \frac{0+1}{6+9} = \frac{1}{15}$$

$$P(\text{lemon} \mid \neg\text{fruit}) = \frac{1+1}{11+9} = \frac{2}{20} = \frac{1}{10}$$

$$P(\text{yellow} \mid \text{fruit}) = \frac{0+1}{6+9} = \frac{1}{15}$$

$$P(\text{yellow} \mid \neg\text{fruit}) = \frac{2+1}{11+9} = \frac{3}{20}$$

$$P(\text{Blue} \mid \text{fruit}) = \frac{0+1}{6+9} = \frac{1}{15}$$

$$P(\text{Blue} \mid \neg\text{fruit}) = \frac{1+1}{11+9} = \frac{2}{20} = \frac{1}{10}$$

$$P(\text{Apricot} \mid \text{fruit}) = \frac{1+1}{6+9} = \frac{2}{15}$$

$$P(\text{Apricot} \mid \neg\text{fruit}) = \frac{0+1}{11+9} = \frac{1}{20}$$

$$P(\text{Apple} \mid \text{fruit}) = \frac{2+1}{6+9} = \frac{3}{15} = \frac{1}{5}$$

$$P(\text{Apple} \mid \neg\text{fruit}) = \frac{0+1}{11+9} = \frac{1}{20}$$

$$P(\text{Banana} \mid \text{fruit}) = \frac{1+1}{6+9} = \frac{2}{15}$$

$$P(\text{Banana} \mid \neg\text{fruit}) = \frac{0+1}{11+9} = \frac{1}{20}$$

$$P(\text{fruit} \mid d6) = P(\text{fruit}) \times P(\text{Orange} \mid \text{fruit}) \times P(\text{Mango} \mid \text{fruit}) \times P(\text{Melon} \mid \text{fruit})$$

$$= \frac{2}{5} \times \frac{2}{15} \times \frac{2}{15} \times \frac{1}{17}$$

$$= 0.000418$$

$$P(\neg\text{fruit} \mid d6) = P(\neg\text{fruit}) \times P(\text{Orange} \mid \neg\text{fruit}) \times P(\text{Mango} \mid \neg\text{fruit}) \times P(\text{Melon} \mid \neg\text{fruit})$$

$$= \frac{3}{5} \times \frac{5}{20} \times \frac{1}{20} \times \frac{1}{22}$$

$$\cancel{= 0.000487}$$

$$= 0.000341$$

d6 → class fruit "~~No~~" Yes

$P(\text{fruit} \mid d7) = P(\text{fruit}) \times P(\text{Orange} \mid \text{fruit})$
$\times P(\text{Red} \mid \text{fruit})$
$\times P(\text{lemon} \mid \text{fruit})$
$\times P(\text{yellow} \mid \text{fruit})$

$$= \frac{2}{5} \times \frac{2}{15} \times \frac{1}{15} \times \frac{1}{15} \times \frac{1}{15}$$

$$= 0.0158 \times 10^{-3}$$

$P(\neg\text{fruit} \mid d7) = P(\neg\text{fruit}) \times P(\text{orange} \mid \neg\text{fruit})$
$\times P(\text{Red} \mid \neg\text{fruit})$
$\times P(\text{lemon} \mid \neg\text{fruit})$
$\times P(\text{yellow} \mid \neg\text{fruit})$

$$= \frac{3}{5} \times \frac{5}{20} \times \frac{3}{20} \times \frac{1}{10} \times \frac{3}{20}$$

$$= \cancel{0.5144 \times 10^{-3}}$$

$$= 0.3375 \times 10^{-3}$$

d7 → class fruit "No"

PROBLEM#2

A)

Concept drift refers to the gradual change over time of the concept underlying a class.
such as US President from Barack Obama to Trump.

The Bernoulli model is particularly robust with respect to concept drift. it can have good performance when the feature size is small. The most important indicator for a class are less likely to change. Thus; a model that only relies on these features is more likely to maintain a certain level of accuracy in concept drift.

Scanned with CamScanner

Answer: 2B:

The NB Classifier makes the conditional independence assumption to reduce the number of parameters. i.e the presence of one term does not affect others. This assumption is very naive for a language model. Thats why it is called naive

Just because the prediction is correct, it does not mean that the estimate and true probability are also same, as the probabilities are directly proportional not equal, but correct estimation does imply accurate prediction

Answer: 2c

If a documents terms do not provide clear evidence for one class versus another, we choose the one that has a higher prior probability.

If one term is not present in the training set, we don't consider its probability 0, as that will reduce all terms to 0. instead we use laplace smoothing to handle unseen term

PROBLEM#3

Taking $Tf = 1 + \log_2(tf_{t,d} + 1)$

$idf = \log_2\left(\frac{N}{Df+1}\right)$

$tf \times idf$

$d_1)$ Orange $\xrightarrow{tf}$ $1 + \log_2(tf_{t,d} + 1)$

$= 1 + \log_2(2+1)$

$= 1 + 1.584$

$= 2.584$

$idf = \log_2\left(\frac{5}{4+1}\right) = 0$

$d_1)$ Lemon $\xrightarrow{tf}$ $1 + \log_2(tf_{t,d} + 1)$

$= 1 + 1$

$= 2$

$idf = \log_2\left(\frac{5}{1+1}\right) = 1.32$

$d_1)$ Red $\xrightarrow{tf}$ $1 + \log_2(tf_{t,d} + 1)$

$= 1 + \log_2(1+1)$

$= 2$

$idf = \log_2\left(\frac{5}{3}\right) = 0.736$

idf:

Blue	Yellow	Apricot	Apple	Mango
$\log_2\left(\frac{5}{3}\right)$	$\log_2\left(\frac{5}{3}\right)$	$\log_2\left(\frac{5}{2}\right)$	$\log_2\left(\frac{5}{3}\right)$	$\log_2\left(\frac{5}{2}\right)$
$= 0.736$	$= 0.736$	$= 1.32$	$= 0.736$	$= 1.32$

Banana

$= 1.32$

	Orange	lemon	Red	Blue	Yellow	Apricot	Apple	Mango	Banana	Melon
d_1 <	0.[illegible]	2.64	1.264	0.736	0.736	1.32	0.736	1.32	1.32	2.32
d_2 <	0.[illegible]	1.32	1.264	1.264	1.264	1.32	0.736	1.32	1.32	2.32
d_3 <	0.[illegible]	1.32	0.736	0.736	0.736	2.64	1.264	2.64	1.32	2.32
d_4 <	0.[illegible]	1.32	0.736	0.736	0.736	1.32	0.736	1.32	2.64	2.32
d_5 <	0.[illegible]	1.32	0.736	0.736	1.264	1.32	0.736	1.32	1.32	2.32
d_6 <	0.[illegible]	1.32	0.736	0.736	0.736	1.32	0.736	2.64	1.32	4.64
d_7 <	0.[illegible]	2.64	1.264	0.736	0.736	1.32	0.736	1.32	1.32	2.32

Melon → $\log_2\left(\frac{5}{0.1}\right) = 2.32$

b)

$el_{fruit\ Yes} = \frac{1}{2} < 0, 2.64, 1.464, 1.464, 1.464, 3.96, 2.2, 3.96, 3.96 >$

$= < 0, 1.32, 0.736, 0.736, 0.736, 1.98, 1.1, 1.98, 1.98 >$

$el_{fruit = No} = \frac{1}{3} < 0, 5.28, 3.664, 2.928, 3.664, 1.32\times3, 0.736\times3, 1.32\times3, 1.32\times3 >$

$= < 0, 1.76, 1.22, 0.976, 1.22, 1.32, 0.736, 1.32, 1.32 >$

c) d_6

$|\vec{el}_{yes} - \vec{v}(d_6)| = \sqrt{0.4356 + 0.132 + 0.4356 + 0.4356 + 5.3824}$

$= 2.611$

$|\vec{el}_{NO} - \vec{v}(d_6)| = \sqrt{0.1936 + 0.234 + 0.0576 + 0.2346 + 0 + 0 + 1.7424 + 0 + 5.3824}$

$= 2.8$

Hence $d_6 \rightarrow$ "yes"

$$|\overrightarrow{\mu_{Yes}} - \vec{v}(d_7)| = \sqrt{1.7424 + 0.0595 + 0.0576 + \cancel{0.2574} + 0.4356 + 0.1324 + 0.4356 + 0.4356}$$

$$= \cancel{1.296}\ 1.9266$$

$$|\overrightarrow{\mu_{No}} - \vec{v}(d_7)| = \sqrt{\cancel{0.7744 + 0.0594} \cancel{0.0576}}$$

$$= \sqrt{(2.64 - 1.76)^2 + (1.464 - 1.22)^2 + (0.976 - 0.736)^2 + (1.22 - 0.732)^2}$$

$$= 1.06$$

then.

$$d_7 \longrightarrow \text{"No"})$$

3(d)

$$\cos(d_1, d_6) = \frac{\vec{d_1} \cdot \vec{d_6}}{|\vec{d_1}||\vec{d_6}|}$$

$$= \frac{2349}{(4.62)(5.99)}$$

$$= 0.864$$

$\cos(d_2, d_6) = 0.882$

$\cos(d_3, d_6) = 0.887$

$\cos(d_4, d_6) = 0.878$

$\cos(d_5, d_6) = 0.922$

3

d_5	0.922
d_3	0.887
d_2	0.882

$d_6 \rightarrow$ "No"

$\cos(d_1, d_7) = 0.999$

$\cos(d_2, d_7) = 0.908$

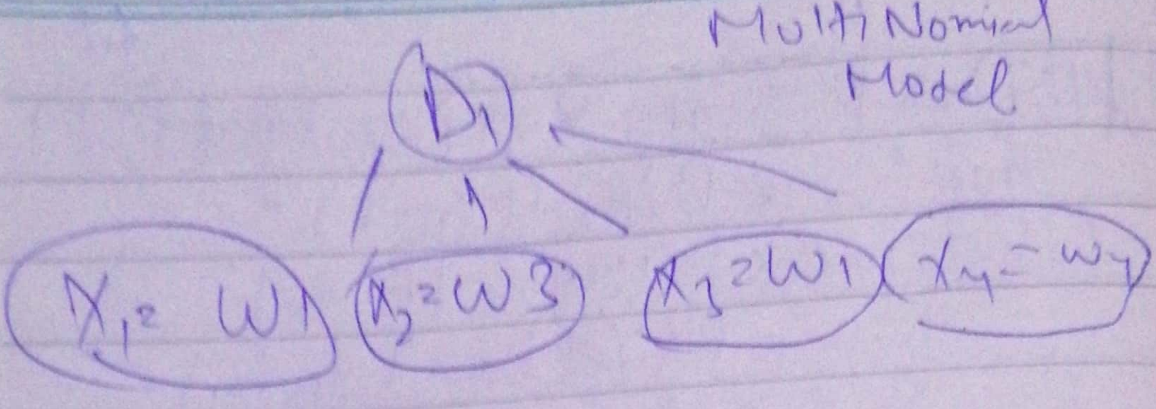
$\cos(d_3, d_7) = 0.839$

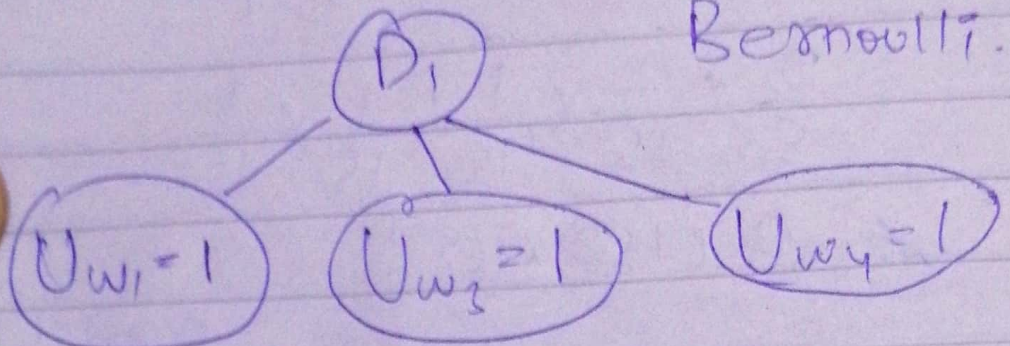
$\cos(d_4, d_7) = 0.8688$

$\cos(d_5, d_7) = 0.912$

$(\cancel{d_6}, d_7)$ $\qquad d_7 \longrightarrow$ "No"

PROBLEM # 2 D

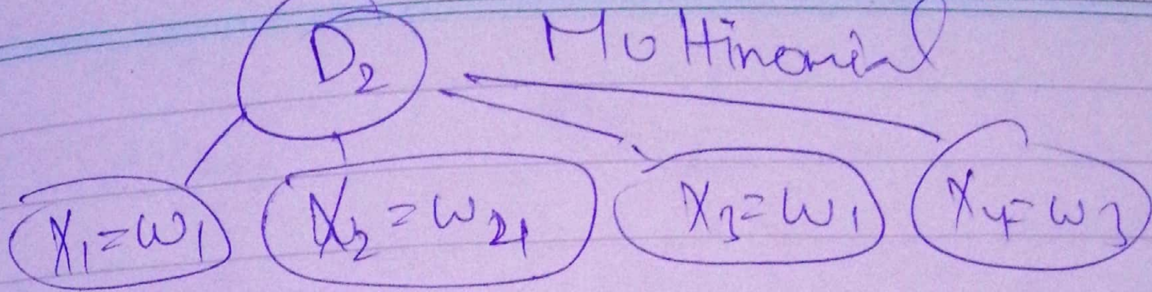


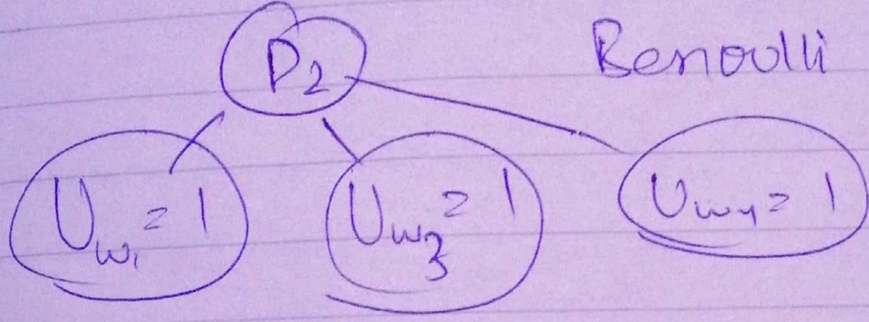


D_1 has different representation as bernoulli representation does not count frequency of features (terms) only if they are present or not.

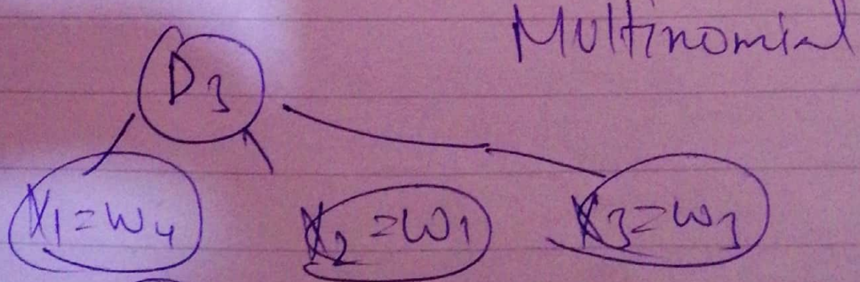
Multinomial Rep is same for D_1 and D_2

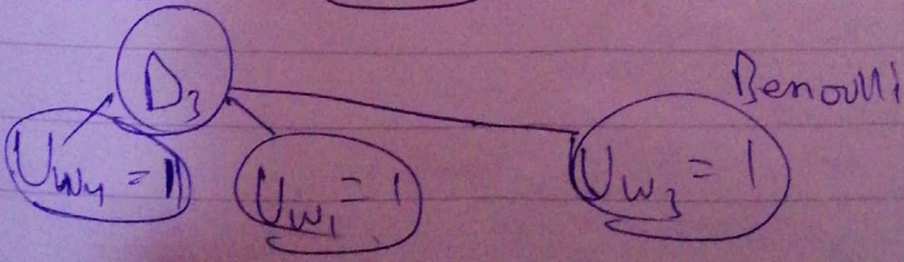
Binomial Rep is same for all D_1, D_2 and D_3





D_2 has different representation as Multinomial representation considers frequency of features (term) while Benoulli only considers presence regardless of count.





D_3 has same representation due to same term count.

Scanned with CamScanner