NLP Sheet 3

1. Find the class naive bayes will assign to "I always like foreign films"

→ wanted d

$$\hat{C} = \underset{c \in C}{\operatorname{argmax}}\ P(d|C)P(C)$$ ← mentioned that it's equal for both (can ignore)

- For C = +ve ← $P(d|C=+ve)$

$$P(\text{"I"}, \text{"always"}, \ldots, \text{"films"} \mid +ve)$$

$$= P(\text{"I"}|+ve)P(\text{"always"}|+ve)P(\text{"like"}|+ve)$$
$$P(\text{"foreign"}|+ve)P(\text{"films"}|+ve)$$

$$= 0.09 \times 0.07 \times 0.29 \times 0.04 \times 0.08$$

$$= 5.8464 \times 10^{-6}$$

- For C = -ve (same way)

$$P(d|C=-ve) = 0.16 \times 0.06 \times 0.06 \times 0.15 \times 0.11$$
$$= 9.504 \times 10^{-6}$$

→ Negative class wins $P(C=-ve|d) > P(C=+ve|d)$

- Predict $\hat{C} = -ve$

- Clearly an incorrect prediction
- 'I', 'foreign', 'films' were given to be more probable for -ve class

2. Given the following short movie reviews and their labels (dataset)

1. Fun, couple, love, love : Comedy

   Fast, Furious, Shoot : action

   couple, Fly, Fast, Fun, Fun : Comedy

   Furious, Shoot, Shoot, Fun : action

   Fly, Fast, Shoot, love : action

d

- Given new review: Fast, couple, Shoot, Fly
  - → Compute most likely class
  - → Use add-1 smoothing

$$\hat{C} = \underset{c \in C}{\operatorname{argmax}} \; P(d|c)P(c)$$

how many times 'Fast' appeared in action documents

	Comedy	action
Fast	1 +1	2 +1
couple	2 +1	0 +1
Shoot	0 +1	4 +1
Fly	1 +1	1 +
Prior	2	3

how many Comedy documents

by counting from dataset
- then +1
- will need to divide by #words in class:

Comedy: 9
action: 11

+ len(vocab) ← for add1

to get probability

Column of Counts

- $P(d|Comedy)P(Comedy) = \frac{1}{(9+7)^4}(2 \cdot 3 \cdot 1 \cdot 2) \cdot \frac{2}{5}$

$= 7.324 \times 10^{-5}$

(i.e. $\frac{2}{9+7} \cdot \frac{3}{9+7} \cdot \frac{1}{9+7} \cdot \frac{2}{9+7} \cdot \frac{2}{5}$)

$P(Past|Comedy)$

$P(Comedy) = \frac{N_c}{N_{doc}}$

- $P(d|action)P(action) = \frac{1}{(11+7)^4}(3 \cdot 1 \cdot 5 \cdot 2) \cdot \frac{3}{5}$

$= 1.7146 \times 10^{-4}$

→ Predict action class

3. Given Dataset

document	'good'	'Poor'	'great'	Class
1	3	0	3	+Ve
2	0	1	2	+Ve
3	1	3	0	-Ve
4	1	5	2	-Ve
5	0	2	0	-Ve

→ Train two models (& test on given Sentence)

- Binarized Naive Bayes
- Multinomial Naive Bayes

1. Multinomial

|V|

	'good'	'Poor'	'great'	#words	Prior
+Ve	3+1	1+1	5+1	9	2/5
-Ve	2+1	10+1	2+1	14	3/5

- Find Probability of "a good, good Plot and great Characters but Poor acting" belonging to each class
  → Circled Known words (unknown ones are ignored)

$$P(d|C=+ve) \overset{P(C)}{=} \frac{1}{(9+3)^4} \cdot (4.4.6.2) \cdot \frac{2}{5}$$

$$= 3.7 \times 10^{-3}$$

$$P(d|C=-ve) \overset{P(C)}{=} \frac{1}{(14+3)^4} (3.3.3.11) \cdot \frac{3}{5}$$

$$= 2.1336 \times 10^{-3}$$

⇒ Predict Positive ✓

2. Binarized

	'good'	'Poor'	'great'	#words	Prior
+ve	1_{+1}	1_{+1}	2_{+1}	4	215
-ve	2_{+1}	3_{+1}	1_{+1}	6	315

- To get the counts, set all those in the given to 1 (equivalent to removing duplicates from each document)

1 0 1 +
0 1 1 +
1 1 0 -
1 1 1 -
0 1 0 -

Can't good only one!

- Repeat ① → remove duplicates from test seq.

$$P(d|C=+ve) \overset{P(C)}{=} \frac{1}{(4+3)^3} (2.3.2) \cdot \frac{2}{5} = 0.014$$

$$P(d|C=-ve)P(C) = \frac{1}{(6+3)^3}(3.2.4).\frac{3}{5} = 0.0197$$

$\Rightarrow$ Predict -ve ✓

* The results from the two models don't agree
  → hard to tell which is more correct (the sentiment is mixed, even for human judgement)
  . if we normalize the probabilities to get $P(C|d)$ in each case we find that multinomial is significantly more confident (65% vs. 58%) because of how good is repeated.
  → binarized would've felt more correct if
  . if the sentence rather was "good characters and good plot but the acting was bad enough to make me hate the movie"

4. Consider the following confusion matrix

model \ Gold	spam	not spam
spam	70	30
not spam	70	330

not spam = ham

$P_{spam} = 70/(70+30) = 0.7$

$P_{not\ spam} = 330/(70+330) = 0.825$

$R_{spam} = \frac{70}{70+70} = 0.5$ $\quad R_{not\ spam} = \ldots$

. if a specific class is given to be the +ve class
  → Precision, recall should be computed for it (unless otherwise mentioned)

- here it directly asked for Precision & recall w.r.t to Spam class

$$P = 0.7 \qquad R = 0.5 \qquad F1 = \frac{2PR}{P+R} = 0.583$$

→ Noting that it has high Precision compared to recall

→ Must imply it has low False Positives compared to False negatives

i.e., Spam emails let into inbox > ham emails that go to Spam

* Consider two users

A: "I hate when I see Spam email in inbox" ← wants low FN

"I'm okay with checking if some ham emails are in Spam periodically" ← okay if FP is high

- Model is clearly not for her. (FN > FP)

✓ B: "what is Spam folder? Spam is okay for me." ← okay if FN is high

"I hate when my ham emails aren't in my inbox" ← wants low FP

- Model is clearly for him.

5. For Logistic regression

→ Show that deciding the Class to classify x as by Checking whether $P(C=1|x) > \frac{1}{2}$ or not is equivalent to deciding by a linear decision boundary in the input space

$$P(y=1|x) > 0.5$$

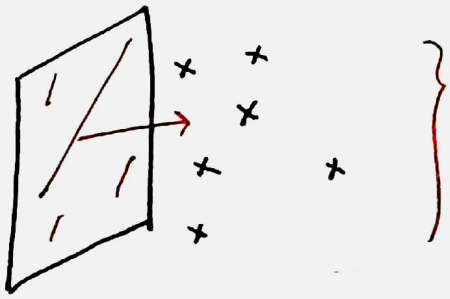
$$\frac{1}{1+e^{-(wx+b)}} > 0.5 \quad \text{. Plug } P(y=1|x)$$

$$1 + e^{-(wx+b)} < 2 \quad \text{. flip both sides}$$

$$e^{-(wx+b)} < 1 \quad \text{. } -1$$

$$-(wx+b) < 0 \quad \text{. take ln of both sides}$$

$$\underbrace{wx+b > 0}_{\text{hyperplane in input space } (x)} \quad \text{. } \times -1$$



} equivalent to the Prob Check is Checking if X is above the hyperplane.

6. → Consider a trained binary (2 classes) logistic model

- Two features x_1 and x_2 and learnt weights $w_1 = 0.2$, $w_2 = 0.4$, $b = 0.5$

i. * Write the Output Prob. equations for this classifier

$$\rightarrow z = 0.2x_1 + 0.4x_2 + 0.5$$

$$P(y=1|x) = \frac{1}{1+e^{-(0.2x_1+0.4x_2+0.5)}} \quad (z)$$

$$P(y=0|x) = \frac{1}{1+e^{z}} = 1 - P(y=1|x)$$

ii) let

$x_1 = 1$ if doc. has '?', else $x_1 = 0$

$x_2 =$ if doc. has 'what', else $x_2 = 0$

- let $y = 1$ (+ve class) be 'doc. is question' and $y = 0$ (-ve class) be 'doc. isn't'

- Find the classification of the following

$z > 0 \leftrightarrow y = 1$

document	true label	→ Model Output
weather is great.	$y=0$	1
what is NLP?	$y=1$	1
what do you think	$y=1$	1

- Compute acc., P, R, F1

model \ gold	+ve	-ve	
+ve	2	1	$P = \frac{2}{3}$
-ve	0	0	
	$R = \frac{2}{2}$		$acc = \frac{2+0}{2+1+0+0} = 2/3$

$$F1 = \frac{2PR}{P+R} = \frac{4/3}{1+\frac{2}{3}} = 0.8$$