

# NLP Sheet 3

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

→ Wanted  $d$

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

mentioned that  
it's equal for both  
(can ignore)

• For  $C = +ve$

$$P(d|C=+ve)$$

$$P("I", "always", \dots, "Films" | +ve)$$

$$= P("I" | +ve) P("always" | +ve) P("like" | +ve) \\ P("Foreign" | +ve) P("Films" | +ve)$$

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

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

• For  $C = -ve$  (Some 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, take : Comedy

Past, Furious, Shoot : action

Couple, Fly, Past, Fun, Fun : Comedy

Furious, Shoot, Shoot, Fun : action

Fly, Past, Shoot, love : action

d

Given new review Past, Couple, Shoot, Fly  
 → Compute most likely class  
 → Use add-1 Smoothing

$$\hat{C} = \arg \max_{C \in \mathcal{C}} P(d|C) P(C)$$

	Comedy	Action	
Past	1 +1	2 +1	how many times 'Past' appeared in action documents
Couple	2 +1	0 +1	
Shoot	0 +1	4 +1	
Fly	1 +1	1 +	} by Counting from dataset
Prior	2	3	+ then +1
	↑ how many Comedy documents		. Will need to divide by # words in class:
			Comedy      Action
			9            11
			Por add 1 + len(vocab) ← to get Probability

Column of Conts

$$\bullet P(\text{d1Comedy})P(\text{Comedy}) = \frac{1}{(9+7)^4} (2.3.1.2) \cdot \frac{2}{5}$$

$$= 7.324 \times 10^{-5} \quad (\text{i.e. } \frac{2}{9+7} \cdot \frac{3}{9+7} \cdot \frac{1}{9+7} \cdot \frac{2}{9+7} \cdot \frac{2}{5})$$

$\downarrow P(\text{Past1Comedy})$

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

$$\bullet P(\text{d1action})P(\text{action}) = \frac{1}{(11+7)^4} (3.1.5.2) \cdot \frac{3}{5}$$

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

$\rightarrow$  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

$\rightarrow$  Train two models (& test on given sentence)

- Binarized Naive Bayes
- Multinomial Naive Bayes

#### 1. Multinomial

	'good'	'Poor'	'great'	#Words	Prior
+ve	$3+1$	$1+1$	$5+1$	9	$2/5$
-ve	$2+1$	$10+1$	$2+1$	14	$3/5$

①

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

$$P(d|C=+ve) = \frac{1}{(9+3)^4} (4.4.6.2) \cdot \frac{2}{5} \\ = 3.7 \times 10^{-3}$$

$$P(d|C=-ve) = \frac{1}{(14+3)^4} (3.3.3.11) \cdot \frac{3}{5} \\ = 2.1336 \times 10^{-3}$$

⇒ Predict Positive

## 2. Binarized

	'good'	'Poor'	'great'	#words	Prior
+ve	1 <sub>+1</sub>	1 <sub>+1</sub>	2 <sub>+1</sub>	4	2/5
-ve	2 <sub>+1</sub>	3 <sub>+1</sub>	1 <sub>+1</sub>	6	3/5

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

Cnt good only once!

1 0 1 \*  
 0 1 1 \*  
 1 1 0 \*  
 0 1 0 \*

- Repeat ① → Remove duplicates from test seq.

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

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

$\Rightarrow$  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.

→ binomial would've felt more correct 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

		Gold		
		spam	not spam	
model	spam	70	30	$P_{\text{spam}} = 70 / (70 + 30) = 0.7$
	not spam	70	330	$P_{\text{not spam}} = 330 / (70 + 330) = 0.825$
		$R_{\text{spam}} = \frac{70}{70+70} = 0.5$	$R_{\text{not spam}} = \dots$	

not spam = ham

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

• Here it directly asked for Precision & recall w.r.t to Spam class

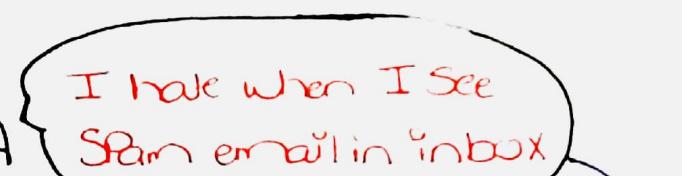
$$P = 0.7 \quad R = 0.5 \quad F_1 = \frac{2PR}{P+R} = 0.583$$

→ Note that it has high precision compared to recall

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

i.e., Spam emails let  $\rightarrow$  ham emails that go to inbox  $\rightarrow$  Spam

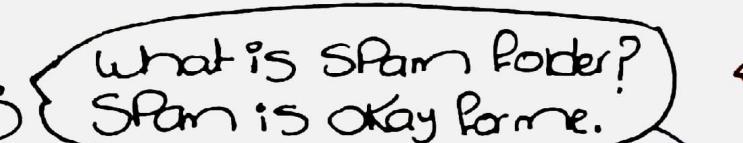
\* Consider two users

A  I hate when I see Spam email in inbox  $\leftarrow$  wants low Fn

I'm okay with checking if some ham emails are in Spam Periodically

$\leftarrow$  Okay if PP is high

. Model is clearly not for her. ( $F_n > P_P$ )

B  what is Spam folder? Spam is okay for me.

$\leftarrow$  Okay if Fn is high

I hate when my ham emails aren't in my inbox

$\leftarrow$  wants low PP

. 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.

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

$$\frac{1}{1 + e^{-\frac{(wx+b)}{z}}} > 0.5$$

• Plug  $P(y=1|x)$

$$1 + e^{-\frac{(wx+b)}{z}} < 2$$

• Flip both sides

$$e^{-\frac{(wx+b)}{z}} < 1$$

•  $\downarrow$

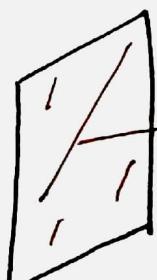
$$-(wx+b) < 0$$

• take  $\ln$  of both sides

$$wx + b > 0$$

$\underbrace{\text{HyperPlane}}$

in input space ( $x$ )



} equivalent to the prob  
Check is checking  
if  $x$  is above the  
hyperPlane.

6. → Consider a trained <sup>2 classes</sup> binary 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

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

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

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

ii) let

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

$x_2 = 1$  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

$$F_1 = \frac{2PR}{P+R} = \frac{4/3}{1+\frac{2}{3}} = 0.8$$

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