Steps in ~~Naïve Bayes~~

Multinomial Naive Bayes Example

Q.

Doc ID	Document	Class
1	Love Happy Joy Joy Happy	Yes
2	Happy Love Kick Joy Happy	Yes
3	Love Move Joy Good	Yes
4	Love Happy Joy Love Pain	Yes
5	Joy Love Pain kick Pain	No
6	Pain Pain Love kick	No
Test 7	Love Pain Joy Love Kick	?

$$\hat{P}(c) = \frac{N_c}{N} \qquad \hat{P}(w|c) = \frac{count(w,c) + 1}{count(c) + |v|}$$

∴ Priors $P\left(\frac{y}{c}\right) = \frac{4}{6} = 0.667$

$P\left(\frac{N}{c}\right) = \frac{2}{6} = 0.833$

Vocabulary : { Love, Happy, Joy, Kick, Move, Good, Pain }

Count term frequency in each class

Term	C(Yes)	C(No)
Love	5	2
Happy	5	0
Joy	5	1
Kick	1	2
Move	1	0
Good	1	0
Pain	1	4
	19	9

$$\therefore P(w|Yes) = \frac{C_{Yes}(w)+1}{19+7}$$

$$\therefore P(w|No) = \frac{C_{No}(w)+1}{9+7}$$

Likelihood of Doc 7

$$P(Yes|d_7) = P($$

working in log space to avoid underflow

$$\log P(Yes|d_7) = \log P(Yes) + 2\log P(Love|Yes) + \log P(Pain|Yes) + \log P(Joy|Yes) + \log P(Kick|Yes)$$

$$= \log\left(\frac{2}{3}\right) + 2\log\left(\frac{5+1}{26}\right) + \log\left(\frac{1+1}{26}\right) + \log\left(\frac{5+1}{26}\right) + \log\left(\frac{1+1}{26}\right)$$

$$= -9.9344$$

$$\log P(No|d_7) = \log P(No) + 2\log P(Love|No) + \log P(Pain|No) + \log P(Joy|No) + \log P(Kick|No)$$

$$= \log\left(\frac{1}{3}\right) + 2\log\left(\frac{2+1}{16}\right) + \log\left(\frac{4+1}{16}\right) + \log\left(\frac{1+1}{16}\right) + \log\left(\frac{2+1}{16}\right)$$

$$= -9.3631$$

$$\therefore P(No|d_7) > P(Yes|d_7)$$

$$\therefore \text{Document is classified as No.}$$