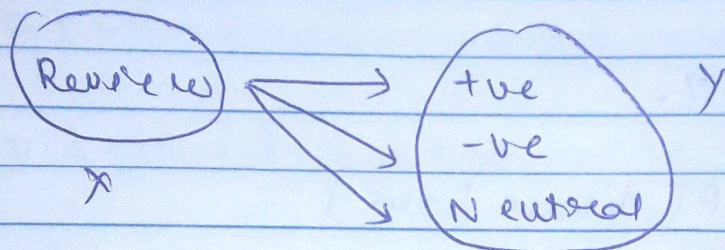
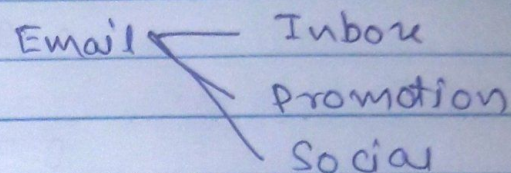
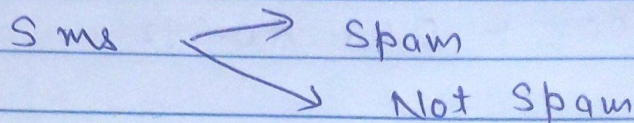


## ⑤ Naive Bayes Classifier:

Text Classify:



Discrete features: 

		cricket		I		like	
0	0	1	0	1	0	0	1

161

Bag of words model

$$P(y_i | x) = \frac{P(x | y_i) P(y_i)}{P(x)}$$

+ve

$$\textcircled{1} P(y=1 | x) = \frac{P(x | y=1) P(y=1)}{P(x)}$$

-ve

$$\textcircled{2} P(y=0 | x) = \frac{P(x | y=0) P(y=0)}{P(x)}$$

$\leftarrow$  Ignoring in both cases

$$\arg \max_i (P(y_i | x))$$

= max of ① and ②

$$= \arg \max_i (P(x | y_i) P(y_i))$$



$$P(y=1) = \frac{\text{Count all +ve Review}}{\text{Total Reviews}}$$

$$= \frac{\sum_{i=1}^m 1 \{y^{(i)} = 1\}}{m} \quad (\text{add 1 when +ve})$$

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

$$x = \langle x_1, x_2, x_3, \dots, x_{|V|} \rangle$$

total no. of vocabs

$$P(x_1, x_2, x_3, \dots, x_{|V|} | y=1)$$

$$= P(x_1 | y=1) \cdot P(x_2 | y=1, x_1) \cdot$$

$$P(x_3 | y=1, x_1, x_2) \cdot \dots$$

$$P(x_{|V|} | y=1, \underbrace{x_1 \dots x_{|V|-1}}_{\text{given thing}})$$

-- good --
-- awesome --
-- like --

$$P(\text{awesome} | +ve, \text{good})$$

$$P(\text{like} | +ve, \text{good, awesome})$$

Computation is not easy that is why

Naive assumption came into picture :



$$p(x_i | y, \underbrace{x_1 \dots x_{i-1}}_{\text{not dependent upon all features}})$$

$$= p(x_i | y)$$

e.g.

$$= p(\text{awesome} | +ve)$$

$$\text{or } p(\text{like} | +ve)$$

∴ Naïve Bayes Assumption

$$\Rightarrow p(x | y=1)$$

$$= p(x_1 | y=1) \cdot p(x_2 | y=1) \cdot \dots \cdot p(x_M | y=1)$$

$$= \prod p(x_i | y=1)$$

$$p(y=1 | x) = \frac{\prod p(x_i | y=1) \cdot p(y=1)}{p(x)}$$

$$p(x) = p(x | y=0) p(y=0) + p(x | y=1) p(y=1)$$

$$p(y=0 | x) = \frac{\prod p(x_i | y=0) p(y=0)}{p(x)}$$

$$= 1 - p(y=1 | x)$$

(for Binary Classification)