(1)

$P(+)$

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

$\Rightarrow \log(0.09) + \log(0.07) + \log(0.29)$

$+ \log(0.04) + \log(0.08)$

$= 5.8 \times 10^{-6}$

$P(-)$

$= 0.16 \times 0.06 \times 0.06 \times 0.15 \times 0.11$

$= 9.5 \times 10^{-6}$

Negative

(2)

$P(com) = \frac{2}{5}$

$P(act) = \frac{3}{5}$

c(com) = 9

c(act) = 11

total = 20

V = 7

fun, couple, love, love → com

fast, furious, shoot act

couple, fly, fast, fun, fun com

furious, shoot, shoot, fun act

fly, fast, shoot, love act

$$P(\text{fast} \mid \text{com})$$

$$= \frac{1+1}{9+7} = \frac{2}{16} = \frac{1}{8}$$

$$P(\text{fast} \mid \text{act})$$

$$= \frac{2+1}{11+7} = \frac{3}{18} = \frac{1}{6}$$

$$P(\text{couple} \mid \text{com})$$

$$= \frac{2+1}{9+7} = \frac{3}{16}$$

$$P(\text{couple} \mid \text{act}) = \frac{1}{11+7} = \frac{1}{18}$$

$$P(\text{shoot} \mid \text{com})$$

$$= \frac{1+1}{9+7} = \frac{1}{8}$$

$$P(\text{shoot} \mid \text{act})$$

$$= \frac{4+1}{11+7} = \frac{5}{18}$$

$$P(\text{fly} \mid \text{com})$$

$$= \frac{1+1}{9+7} = \frac{1}{8}$$

$$P(\text{fly} \mid \text{act})$$

$$= \frac{1+1}{11+7} = \frac{1}{9}$$

$P(com)$

$= \frac{2}{5} \times \frac{2 \times 2 \times 2 \times 3}{16^4} = 1.46 \times 10^{-4}$

$P(act)$

$= \frac{3}{5} \times \frac{3 \times 1 \times 5 \times 2}{18^4} = 1.7 \times 10^{-4}$

$= 0.6 \times$

Action

(3) Because of its time efficiency, and it's good when the dataset doesn't have a lot of data.

(4) Naive assumption stands for each class, they are independent with each other. and they will not consider correlation between them, In sentiment analysis, they only consider each word has its sentiment and not considering the correlation between words. It's still used because it's still performing well under this assumption.