

Page No. \_\_\_\_\_  
Date \_\_\_\_\_

Name : Atharva Virkar 202318004 Bigdata

Cat Document

- just Plain boring
  - entirely predictable and lacks energy
  - no superior and very few laughs
  - + very powerful
  - + the most fun film of the summer.
- } Train

? predictable with no fun Test

1) prior from trains  $p(c_i) = \frac{N_{c_i}}{\sum N} \Rightarrow P(-) = 3/5$   
 $P(+) = 2/5$

2) Likelihood from Training :

$$P(w_i/c) = \frac{\text{count}(w_i, c)}{\sum_{w \in V} \text{count}(w, c) + |V|} \quad \forall U \text{ is 20 unique words}$$

$$P(\text{Predictable}/+) = \frac{0+1}{9+20} = \frac{1}{29}$$

$$P(\text{Predictable}/-) = \frac{1+1}{14+20} = \frac{2}{34} = \frac{1}{17}$$

$$P(\text{fun}/+) = \frac{1+1}{9+20} = \frac{2}{29}$$

$$P(\text{Fun}|-) = \frac{0+1}{14+20} = \frac{1}{34}$$

$$P(\text{no} |+) = \frac{0+1}{9+20} = \frac{1}{29}$$

$$P(\text{no}|-) = \frac{1+1}{14+20} = \frac{1}{12}$$

3) Solovis test set

$$\begin{aligned} \rightarrow P(-) \cdot P(\text{pred}|-) \cdot P(\text{No}|-) \cdot P(\text{fun}|-) \\ \approx \underline{\underline{6.106 \times 10^{-5}}} \end{aligned}$$

$$\begin{aligned} \rightarrow P(+) \cdot P(\text{predictable} |+) \cdot P(\text{No}+ ) \cdot P(\text{fun} |+) \\ \approx \underline{\underline{3.2 \times 10^{-5}}} \end{aligned}$$