

BIG DATA

Henry Macwan

Sid :- 202918060

Training

cat

Documents

- just plain boring
- entirely predictable & lacks energy
- no surprises & very few laughs
- + very powerful
- + The most fun film of summer

Test

?

predictable with no fun.

Soln

Priors :

$$P(-) = 3/5 \quad \& \quad P(+) = 2/5$$

conditional probabilities :-

$$P(\text{predictable} | -) = \frac{1+1}{14+20} = \frac{1}{17}$$

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

$$P(\text{no} | -) = \frac{1+1}{14+20} = \frac{1}{17}$$

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

$$P(\text{predictable} | +) = \frac{0+1}{9+24} = \frac{1}{33}$$

$$P(\text{with} | +) = \frac{0+1}{9+24} = \frac{1}{33}$$

$$P(\text{no} | +) = \frac{0+1}{9+24} = \frac{1}{33}$$

$$P(\text{fun} | +) = \frac{1+1}{9+24} = \frac{2}{33}$$

choosing the category :-

$$P(+ | \text{test}) = P(\text{predictable} | +) \times P(\text{with} | +) \times P(\text{no} | +) \times P(\text{fun} | +) \times P(+)$$

$$= \frac{1}{33} \times \frac{1}{33} \times \frac{1}{33} \times \frac{2}{33} \times \frac{2}{5} = \frac{4}{5929605}$$

$$= 6.7459119e^{-7} \approx 0.0000007$$

$$P(- | \text{test})$$

$$= P(\text{predictable} | -) \times P(\text{with} | -) \times P(\text{no} | -) \times P(\text{fun} | -) \times P(-)$$

$$= \frac{1}{17} \times \frac{1}{34} \times \frac{1}{17} \times \frac{1}{34} \times \frac{3}{5}$$

$$= \frac{3}{1670420}$$

$$= 1.79595551e^{-6}$$

$$\approx 0.000002$$

(continued above)