

Cat

Documents

Training

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

Test

? predictable with no fun

Question

Find ? i.e the category of the above test document.

$$P(+)=2/5$$

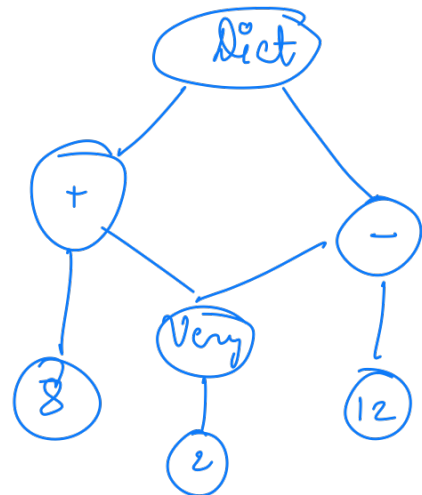
$$P(-)=3/5$$

$$P(\text{predictable})=1/23$$

$$P(\text{with})=0$$

$$P(\text{no})=1/23$$

$$P(\text{fun})=1/23$$



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

$$P(\text{predictable}/+) = \frac{0+1}{29} = 1/29$$

$$P(\text{no}/-) = \frac{1+1}{34} = 1/17$$

$$P(\text{no}/+) = \frac{0+1}{29} = 1/29$$

$$P(\text{fun}/-) = \frac{1}{34}$$

$$P(\text{fun}/+) = 2/29$$

Scoring

$$P(-) P(D/-) = 3/5 \times \frac{2 \times 2}{34^3}$$

$$\approx 6 \times 10^{-5}$$

$$P(+) P(D/+) = \frac{2}{5} \times \frac{2}{29^3}$$

$$\approx 3.2 \times 10^{-5}$$

$$\text{As, } P(-) P(D/-) > P(+) P(D/+)$$

\Rightarrow test document belongs to the negative category.