__/__/____

Asma Nasmawala
20231B025

Example:—

	Category	Documents
Training	−	just plain boring
	−	entirely predictable & lacks energy
	−	no surprises & very few laughs
	+	very powerful
	+	the most fun film of the summer
Test	?	predictable with no fun

⇒ The prior prior for the classes is computed using :- $P(c) = \frac{N_c}{N_{total}}$

$\therefore P(-) = \frac{3}{5}$ and $P(+) = \frac{2}{5}$

⇒ Now, the word 'with' doesn't occur in the training set. Hence we can drop it.

⇒ Let's calculate likelihoods of the remaining 3 words using:-

$$P(w_i|c) = \frac{\text{count}(w_i, c) + 1}{\sum_{w \in V} \text{count}(w, c)) + |V|}$$

$\Rightarrow P('Predictable'|-) = \frac{1+1}{14+20} = \frac{1}{17}$

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

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

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

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

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

$\Rightarrow$ Now, for the test document, class can be computed as follows:-

$P(-)\,P(S|-) = \frac{3}{5} \times \frac{(2)(2)(1)}{34^3} \approx \underline{0.00006}$

$P(+)\,P(S|+) = \frac{2}{5} \times \frac{(1)(1)(2)}{29^3} \approx \underline{0.000032}$

Thus, we can say that the test document is of negative class.