

Headline	Category
"Team wins the final match"	Sports
"Government passes new law"	Politics
"Star player scores a goal"	Sports
"President announces new policy"	Politics
"Coach praises team performance"	Sports
"Senate debates healthcare bill"	Politics

"Team scores in the match"

From all training data, list all unique words:

**["team", "wins", "the", "final", "match",
"government", "passes", "new", "law", "star",
"player", "scores", "a", "goal", "president",
"announces", "policy", "coach", "praises",
"performance", "senate", "debates", "healthcare",
"bill"]**

→ Total unique words (V) = 24

Sports Headlines (Total words = 14):

- team: 2
- wins: 1
- the: 1
- final: 1
- match: 1
- star: 1
- player: 1
- scores: 1
- a: 1
- goal: 1
- coach: 1
- praises: 1
- performance: 1

Politics Headlines (Total words = 12):

- government: 1
- passes: 1
- new: 2
- law: 1
- president: 1
- announces: 1
- policy: 1
- senate: 1
- debates: 1
- healthcare: 1
- bill: 1

$$P(\text{word}|\text{class}) =$$

$$(\text{word count}+1) / \text{total words in class} + V$$

□ **Calculate for Sports: (total words = 14, V = 24)**

. $P(\text{team} | \text{Sports}) = (2+1)/(14+24) = 3/38$

. $P(\text{scores} | \text{Sports}) = (1+1)/(14+24) = 2/38$

. $P(\text{in} | \text{Sports}) = (0+1)/(14+24) = 1/38$

. $P(\text{the} | \text{Sports}) = (2+1)/(14+24) = 3/38$

. $P(\text{match} | \text{Sports}) = (1+1)/(14+24) = 2/38$

Multiply them: (4.54×10^{-7})

$$P(\text{Sports}) \propto 3/38 \times 2/38 \times 1/38 \times 3/38 \times 2/38$$

□ **Calculate for Politics: (total words = 12, V = 24)**

- . $P(\text{team} \mid \text{Politics}) = (0+1)/(12+24) = 1/36$
- . $P(\text{scores} \mid \text{Politics}) = 1/36$
- . $P(\text{in} \mid \text{Politics}) = 1/36$
- . $P(\text{the} \mid \text{Politics}) = 1/36$
- . $P(\text{match} \mid \text{Politics}) = 1/36$

Multiply them: (1.65×10^{-8})

$$P(\text{Politics}) \propto 1/36 \times 1/36 \times 1/36 \times 1/36 \times 1/36$$

$$P(\text{Sports}) > P(\text{Politics})$$