Task A

$P(\text{spam} \mid \text{"cool bicycles and motorbikes"}) = P(\text{spam}) \cdot P(\text{cool} \mid \text{spam}) \cdot P(\text{bicycles} \mid \text{spam}) \cdot P(\text{and} \mid \text{spam}) \cdot P(\text{motorbikes} \mid \text{spam})$

$P(\text{not spam} \mid \text{"cool bicycles and motorbikes"}) = P(\text{not spam}) \cdot P(\text{cool} \mid \text{not spam}) \cdot P(\text{bicycles} \mid \text{not spam}) \cdot P(\text{and} \mid \text{not spam}) \cdot P(\text{motorbikes} \mid \text{not spam})$

$P(\text{spam}) = \frac{3}{6} = \frac{1}{2}$

$P(\text{not spam}) = \frac{3}{6} = \frac{1}{2}$

Laplace smoothing

$$P(w \mid c) = \frac{\text{count}(w \text{ in } c) + 1}{\text{total words in class} + \text{Vocabulary size}}$$

Total words in spam = 14 Vocabulary size = 27

Total words in normal = 25

$P(\text{cool} \mid \text{spam}) = \frac{0+1}{14+27}$ $P(\text{and} \mid \text{spam}) = \frac{2+1}{14+27}$

$P(\text{Bicycles} \mid \text{spam}) = \frac{2+1}{14+27}$ $P(\text{motorbikes} \mid \text{spam}) = \frac{2+1}{14+27}$

$P(\text{cool} \mid \text{normal}) = \frac{1+1}{25+27}$ $P(\text{and} \mid \text{normal}) = \frac{2+1}{25+27}$

$P(\text{bicycles} \mid \text{normal}) = \frac{2+1}{25+27}$ $P(\text{motorbikes} \mid \text{normal}) = \frac{2+1}{25+27}$

$P(\text{spam} \mid \text{message}) = \frac{1}{2} \cdot \frac{1 \cdot 3 \cdot 3 \cdot 3}{41^4} = 4.77 \cdot 10^{-6}$

$P(\text{normal} \mid \text{message}) = \frac{1}{2} \cdot \frac{2 \cdot 5 \cdot 5 \cdot 3}{52^4} = 3.69 \cdot 10^{-6}$

Zhelru Kanat

	sport	politics	culture
classified as sport	200	20	30
classified as politics	10	150	15
classified as culture	60	90	10

$$Precision = \frac{TP}{TP+FP}, \quad Recall = \frac{TP}{TP+FN}, \quad F1 = \frac{2 \cdot Pre \cdot Rec}{Pre + Rec}$$

Class: sport

$TP = 200$

$FP = 20 + 30 = 50$

$FN = 10 + 60 = 70$

$$Pre = \frac{200}{200+50} = 0,8$$

$$Rec = \frac{200}{200+70} = 0.741$$

$$F1 = \frac{2 \cdot 0,8 \cdot 0,741}{0,8 + 0,741} = 0,769$$

Class: politics

$TP = 150$

$FP = 10 + 15 = 25$

$FN = 20 + 90 = 110$

$$Pre = \frac{150}{150+25} = 0.857$$

$$Rec = \frac{150}{150+110} = 0.577$$

$$F1 = \frac{2 \cdot 0.857 \cdot 0.577}{0.857 + 0.577} = 0,689$$

Class culture

$TP = 10$

$FP = 60 + 90 = 150$

$FN = 30 + 15 = 45$

$$Pre = \frac{10}{10+150} = 0.063$$

$$Rec = \frac{10}{10+45} = 0.182$$

$$F1 = \frac{2 \cdot 0.063 \cdot 0.182}{0,063 + 0.182} = 0.093$$

Zhetru Kanat