

# Intro to NLPs

Date \_\_\_\_\_

## Activity Sheet - 7

(Solution)

Q3) TP  $\rightarrow$  [the, cat, on, the, cat]  $\rightarrow$  5

FP  $\rightarrow$  [sat, read]  $\rightarrow$  2

FN  $\rightarrow$  [sitting, a, is, angry]  $\rightarrow$  4

$$\text{Precision} = \frac{5}{5+2} = 0.714$$

$$\text{Recall} = \frac{5}{5+4} = 0.556$$

$$F1 = \frac{2 \times 0.714 \times 0.556}{0.714 + 0.556} = 0.625$$

Q4) Matching bigrams  $\rightarrow$  [the cat, sat on]  $\rightarrow$  2

Total bigrams  $\rightarrow$  6

$$p_1 = \frac{2}{6} = 0.333$$

$$BP = \begin{cases} 1 & \text{if } c \geq \gamma \\ e^{(1-\frac{\gamma}{c})} & \text{if } c < \gamma \end{cases}$$

$c$  = length of generated sentence = 7  
 $\gamma$  = length of reference = 9

$c < \gamma$ , so;

$$BP = e^{(1-\frac{9}{7})} = e^{-2/7} \approx 0.7408$$

Now for BLEU-2:

Matching words  $\rightarrow$  5

Total words  $\rightarrow$  7

$$p_1 = \frac{5}{7} = 0.714$$

$$\log(p_1) = \log(0.714) = -0.336$$

$$\log(p_2) = \log(0.333) = -1.099$$

$$\frac{1}{2} (\log p_1 + \log p_2) = \frac{1}{2} (-0.336 - 1.099) = -0.718$$



$$e^{-0.718} \approx 0.488$$

$$\text{BLEU score} = 0.89 \times 0.488 = 0.431$$

Q5) We need to compute dot product for each pair of reference and candidate embedding.

For "the":

$$\text{the} + \text{the} = (0.1 \times 0.2) + (0.2 \times 0.3) = 0.08$$

$$\text{the} + \text{cat} = (0.1 \times 0.1) + (0.2 \times 0.9) = 0.19$$

$$\text{the} + \text{meowed} = (0.1 \times 0.7) + (0.2 \times 0.1) = 0.09$$

For "dog":

$$\text{dog} + \text{the} = (0.9 \times 0.2) + (0.1 \times 0.3) = 0.21$$

$$\text{dog} + \text{cat} = (0.9 \times 0.1) + (0.1 \times 0.9) = 0.18$$

$$\text{dog} + \text{meowed} = (0.9 \times 0.7) + (0.1 \times 0.1) = 0.64$$

For "barked":

$$\text{barked} + \text{the} = (0.8 \times 0.2) + (0.9 \times 0.3) = 0.43$$

$$\text{barked} + \text{cat} = (0.8 \times 0.1) + (0.9 \times 0.9) = 0.89$$

$$\text{barked} + \text{meowed} = (0.8 \times 0.7) + (0.9 \times 0.1) = 0.65$$

Recall:

$$R_{\text{BERT}} = \frac{1}{3} (0.19 + 0.64 + 0.89) = 0.575$$

For candidate:

$$\text{the} + \text{the} = (0.1 \times 0.2) + (0.3 \times 0.2) = 0.08$$

$$\text{the} + \text{dog} = (0.2 \times 0.9) + (0.3 \times 0.1) = 0.21$$

$$\text{the} + \text{barked} = (0.2 \times 0.8) + (0.3 \times 0.9) = 0.43$$

$$\text{cat} + \text{the} = (0.1 \times 0.1) + (0.9 \times 0.2) = 0.19$$

$$\text{cat} + \text{dog} = 0.18$$



Date \_\_\_\_\_

$$\text{cat} + \text{barked} = 0.89$$

$$\text{meowed} + \text{the} = (0.7 \times 0.1) + (0.1 \times 0.2) = 0.09$$

$$\text{meowed} + \text{dog} = (0.7 \times 0.9) + (0.1 \times 0.1) = 0.64$$

$$\text{meowed} + \text{barked} = (0.7 \times 0.8) + (0.1 \times 0.7) = 0.65$$

$$F_{\text{BERT}} = \frac{2 \times P_{\text{BERT}} \times R_{\text{BERT}}}{P_{\text{BERT}} + R_{\text{BERT}}}$$

$$= \frac{2 \times 0.659 + 0.573}{0.659 + 0.573}$$

$$F_{\text{BERT}} = 0.615$$