

BLEU Score

$$P_n = \frac{\sum_{n\text{-gram} \in C} \min \left(\max_{i=1, \dots, K} \text{Count}_{r_i}(n\text{-gram}), \text{Count}_c(n\text{-gram}) \right)}{\sum_{n\text{-gram} \in C} \text{Count}_c(n\text{-gram})}$$

- (i)
- | | |
|---|---|
| r_1 : love can always find a way | 6 |
| r_2 : love makes anything possible | 4 |
| c_1 : the love can always do | 5 |
| c_2 : love can make anything possible | 5 |

For c_1 :

1-gram:

the denominator $\sum_{n\text{-gram} \in c_1} \text{Count}_c(n\text{-gram}) = 5$

the numerator: $0 + 1 + 1 + 1 + 0 = 3$

so $P_1 = \frac{3}{5}$

2-gram:

the denominator $\sum_{n\text{-gram} \in c_1} \text{Count}_c(n\text{-gram}) = 4$

the numerator: $0 + 1 + 1 + 0 = 2$

so $P_2 = \frac{1}{2}$

BP:

$c = 5$ $r^* = 4$ so $BP = 1$

BLEU: (with $\log = \log_{10}$)

$$\text{BLEU} = BP \times \exp(0.5 \log P_1 + 0.5 \log P_2) = 0.77$$

For C_2 :

1-gram: the denominator: $\sum_{ngram \in C_2} Count_c(gram) = 5$

the numerator: $1 + 1 + 0 + 1 + 1 = 4$

$$\text{So } p_1 = \frac{4}{5}$$

2-gram: the denominator: $\sum_{ngram \in C_2} Count_c(gram) = 4$

the numerator: $1 + 0 + 0 + 1 = 2$

$$\text{So } p_2 = \frac{1}{2}$$

BP:

$$c = 5 \quad r^* = 4 \quad \text{So } BP = 1$$

BLEU:

$$BLEU = BP \times \exp(0.5 \cdot \log \frac{4}{5} + 0.5 \log \frac{1}{2}) = 0.82$$

So According to BLEU Score, C_2 is a better translation.

But personally, I think C_1 is more elegant, native and powerful.

(ii)

r_1 : love can always find a way 6

c_1 : the love can always do 5

c_2 : love can make anything possible 5

For C_1 :

1-gram: $\frac{3}{5}$

2-gram: $\frac{2}{4}$

$$BP: c = 5 \quad r^* = 6, \quad BP = \exp(1 - \frac{r^*}{c}) = 0.82$$

$$BLEU: BP \times \exp(0.5 \log \frac{3}{5} + 0.5 \log \frac{2}{4}) = 0.63$$

For C2:

$$1\text{-gram: } \frac{2}{6}$$

$$2\text{-gram: } \frac{1}{5}$$

$$BP: 0.82$$

$$BLEU: BP \times \exp(0.5 \log \frac{2}{6} + 0.5 \log \frac{1}{5}) = 0.45$$

This time, according to BLEU, C₁ is a better translation.
I do agree that C₁ is a nice translation.

(iii)

This may force the NMT to translate to the **style** of that reference sentence, particularly if the translation reference all come from a single person.

(iv)

advantages:

- Fast and efficient
- more objective, can be measured and compare

Disadvantage:

- may lose subtlety
- tends to be biased towards the reference's style.