



Deep Learning

Dr. Mehran Safayani

safayani@iut.ac.ir

safayani.iut.ac.ir



 <https://www.aparat.com/mehran.safayani>

 https://github.com/safayani/deep_learning_course

Evaluating machine translation

Evaluating machine translation

- French: Le chat est sur le tapis

- Ref1: the cat is on the mat

Ref_count=2

- Ref2: there is a cat on the mat

Ref_count=1

- MT output: the the the the the the the

Count_clip=min(max_ref_count,count)

- Precision= 7/7

modified precision =2/7

	Count	Max ref count	Count clip
the	7	2	2

Evaluating machine translation

- Ref1: the cat is on the mat
- Ref2: there is a cat on the mat
- MT output: the cat the cat on the mat

the	cat	On	mat
2	1	1	1
1	1	1	1

Count_clip=min(max_ref_count,count)

- precision =5/7

unigram	Count	Max ref count	Count clip
the	3	2	2
cat	2	1	1
on	1	1	1
mat	1	1	1
	7		5

Evaluating machine translation

- Ref1: the cat is on the mat
- Ref2: there is a cat on the mat
- MT output: the cat the cat on the mat
- precision = 4/6

The cat	Cat the	Cat On	On the	The mat
1	0	0	1	1
0	0	1	1	1

Count_clip=min(max_ref_count,count)

bigram	Count	Max ref count	Count clip
The cat	2	1	1
Cat the	1	0	0
Cat on	1	1	1
On the	1	1	1
The mat	1	1	1
	6		4

Bleu_score

$$P_n = \frac{\sum_{ngram \in \hat{y}} count_clip(ngram)}{\sum_{ngram \in \hat{y}} count(ngram)}$$

$$Blue_{score} = BP * \exp\left(\frac{1}{\#ngram} \sum_{n=1}^{\#ngram} P_n\right)$$

Brevity Penalty

$$BP = \begin{cases} 1, & MT \text{ Output Length} > \text{reference Output Length} \\ \exp\left(1 - \frac{\text{reference Output Length}}{MT \text{ Output Length}}\right), & \text{otherwise} \end{cases}$$