

# Sequence to sequence models

Bleu score (optional)

## Evaluating machine translation

French: Le chat est sur le tapis.

Reference 1: The cat is on the wat.

Reference 2: There is a cat on the mat.

MT output: the the the the the the.

Precision: Modified precision:

Bleu mbestudy

### Bleu score on bigrams

Example: Reference 1: The cat is on the mat.

Reference 2: There is a cat on the mat.

MT output: The cat the cat on the mat.

	Count	Coutcip	
the cat	2 ←	1	
cat the	( <		et
cat on	( <	( -	
on the	( ←	1 6	
the mat	<b>←</b>	( 6	

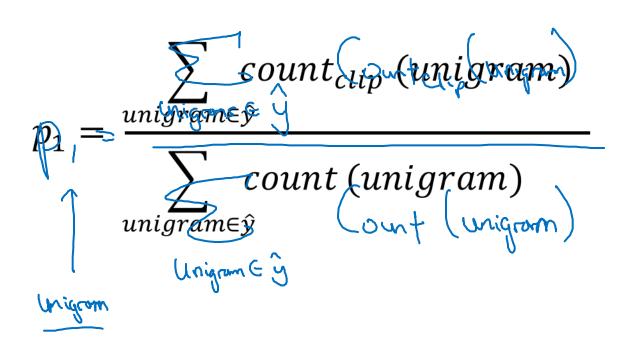
[Papineni et. al., 2002. Bleu: A method for automatic evaluation of machine translation]

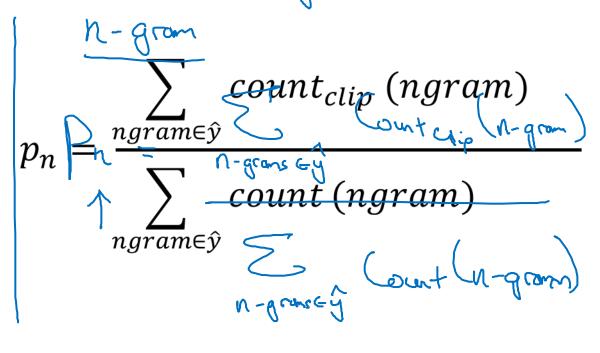
### Bleu score on unigrams

Example: Reference 1: The cat is on the mat.

Reference 2: There is a cat on the mat.

-> MT output: The cat the cat on the mat.





#### Bleu details

$$p_n =$$
Bleu score on n-grams only

P1, P2, P3, P4

[Papineni et. al., 2002. Bleu: A method for automatic evaluation of machine translation]

Andrew Ng