

# Transformer as Machine Translation Evaluation Metrics

Krzysztof Wróbel (Enelpol, Jagiellonian University, AGH University of Science and Technology)

Enelpol



# Data

- use nonblind data in the blind task to increase training data size
- additional corpus:
  - English and Polish subtitles from TED Talks
  - over 80 thousand sentences
  - machine-translated and scored by BLEU

# Methods: BLEU evaluation

- evaluation on nonblind/dev dataset
- BLEU with different smoothing methods from NLTK
  - the best used for scoring TED Talks corpus

# Methods: model

- text regression
- transformer models
- for nonblind: human and machine translated sentences concatenated
- hidden state of the [CLS] token and dense layer with linear activation

# Experiments

- HerBERT large for Polish texts
- XLM-RoBERTa large for mix of English and Polish texts

# Blind

- HerBERT large trained on Ted Talks corpus
- further finetuned on blind/dev, nonblind/dev, and nonblind/test-A with validation on blind/test-A

# Nonblind

- training in different combinations of English sentences, machine-translated and human-translated sentences

# Results: BLEU

The best BLEU smoothing method for all datasets is the second method, which adds 1 to both numerator and denominator.

Method	dev	test-A	test-B
BLEU 0	0.37	0.31	0.32
BLEU 1	0.48	0.41	0.41
BLEU 2	<b>0.59</b>	<b>0.50</b>	<b>0.51</b>
BLEU 3	0.53	0.45	0.46
BLEU 4	0.48	0.41	0.42
BLEU 5	0.46	0.41	0.40
BLEU 7	0.48	0.43	0.42
GLEU	0.53	0.46	0.45



# Submissions

- Baseline 1 - BLEU
- B1 - output from a blind model
- M1/M2 - average from a few models

## Results: Blind

---

	test-B Pearson
M1	0.5076
B1	<b>0.5138</b>
Darek Kłeczek	0.4840
Artur Nowakowski	0.4793

---

## Results: Nonblind

	test-B Pearson
Baseline 1 (BLEU)	0.5103
B1	0.5616
M2	0.5676
M1	0.5677
Artur Nowakowski	0.5728
Darek Kłeczek	<b>0.6137</b>

Source code and models are available at:

<https://github.com/enelpol/poleval2021-task2>

Contact:

[contact@enelpol.com](mailto:contact@enelpol.com)

<https://www.linkedin.com/in/wrobelkrzysztof/>