FORMALITY STYLE TRANSFORMATION (FST)

Team #1: Duc Manh Vu, Suraj Giri, Muslimbek Abduvaliev, Nijat Sadikhov, Akmalkhon Khashimov
Introduction to NLP, University of Bonn

Overview

We experimented with the RNN models and the CoEdit model from Grammarly, transforming an input sentence to a more formal style.By applying various metrics to our models' implementation, we measured the change in formality. Metrics used: a statistical metric of perceived formality, BLEU, TER(p), PINC.

What is formality style transfer?

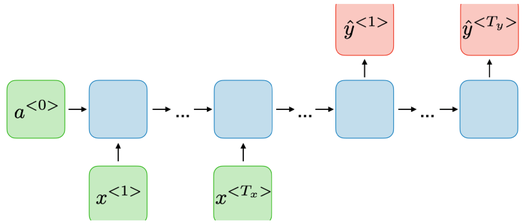
- **Formality style transfer:** A text conversion process switching between informal and formal.
- **The Role of Technology in Learning**: Automatic Formality Style Transfer.
- **Advancements of NLP in Formality Transformation**: Cutting-edge methods (Neural Machine Translation).

Why is it important?

- **Academia:** Maintain formality of academic materials.
- **Business and Service:** Enhance professionalism.
- **Language Learning:** Support improvement of formal writing style.

Many-to-many RNN model

PyTorch



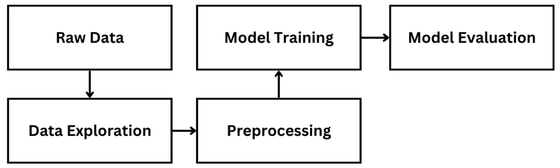
Evaluation:

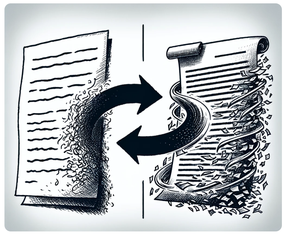
Metric / Set	BLEU Higher=Better	TER+ (TERp) Lower=Better	TER Lower=Better	PINC(n=2) Lower=Better
Validation	0.057	0.943	5.78	0.653
Test	0.061	0.938	5.56	0.639

Dataset: GYAFC (Grammarly's Yahoo Answers Formality Corpus)

- Formality prediction dataset.
- Diverse sources, diverse styles.
- Two main domains:
  Family & Relationships and **Entertainement & Music**.
- One-to-one mapping in the train set.
- Validation set and Test set: four references for each .

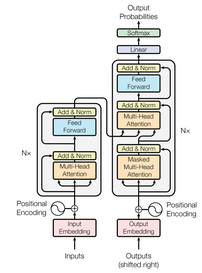
Overall process





CoEdit large

grammarly



Example Input	Example Output
Fix the grammar: When I grow up, I start to understand what he said is quite right.	When I grew up, I started to understand what he said was quite right.
Make this text coherent: Their flight is weak. They run quickly through the tree canopy.	Their flight is weak, but they run quickly through the tree canopy.
Rewrite to make this easier to understand: A storm surge is what forecasters consider a hurricane's most treacherous aspect.	A storm surge is considered a hurricane's most dangerous aspect.
Paraphrase this: They wouldnt want u stepping in.	They would not desire your interference.
Write this more formally: omg i love that song im listening to it right now	I love that song and I am listening to it at this moment.
Write in a more neutral way: The authors' exposé on nutrition studies.	The authors' statements on nutrition studies.

Evaluation:

Pretrained

		BLEU-1 ↑	BLEU-2 ↑	BLEU-3 ↑	BLEU-4 ↑	SacreBLEU ↑	TER ↓	TERp ↓	PINC ↑	Formality ↑
Raw	Val	82.74	74.09	66.66	59.85	59.91	0.3339	0.0845	0.5746	-0.0565
	Test	82.01	73.63	66.38	59.72	59.77	0.33	0.0908	**0.5718**	**-0.0517**
Rule-based preprocess	Val	**83.09**	**74.45**	**66.97**	**60.08**	**60.14**	**0.3316**	**0.0837**	**0.5776**	**-0.0495**
	Test	**82.12**	**73.79**	**66.56**	**58.93**	**59.97**	**0.3289**	**0.09**	0.5692	-0.0523

Finetuned test

BLEU-1 ↑	BLEU-2 ↑	BLEU-3 ↑	BLEU-4 ↑	SacreBLEU ↑	TER ↓	TERp ↓	PINC ↑	Formality ↑
2.33	0.67	0	0	0	4.6043	0.0042	0.9985	0.1616

Metrics for Style Transfer Evaluation

- **Formality:** Scoring formality of n-grams in sentences
- **Bilingual Evaluation Understudy (BLEU):** n-gram matching precision and similarity.
- **Translation Edit Rate (TER):** Measuring edit frequency.
- **Translation Edit Rate Plus (TERp):** Extension of TER, Correlations among stems, synonyms, chosen phases.
- **Paraphrase In N-gram Changes (PINC):** Measures paraphrasing by comparing the ratio of new n-grams to reference n-grams.

Conclusions

- **Our RNN and LSTM-based model:**
  - BiLSTM model.
  - Yielded tragic outcome.
- **CoEdit-large model**:
  - Pretrained CoEdit-large: better results compared to our LSTM-based.
  - Finetuned CoEdit-large: severely deteriorated the result.
- **Future Work - Fluency and Meaning Preservation Metrics:**
  - Explicit Fluency and Meaning preservation evaluation.
  - Optimize grammatical correctness and semantic content retention.

References

- "Dear Sir or Madam, May I Introduce the GYAFC Dataset: Corpus, Benchmarks and Metrics for Formality Style Transfer" (Paper)
- https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks
- https://pytorch.org/docs/stable/generated/torch.nn.LSTM.html
- Vipul Raheja, Dhruv Kumar, Ryan Koo, and Dongyeop Kang. 2023. CoEdIT: Text Editing by Task-Specific Instruction Tuning. In Findings of the Association for Computational Linguistics: EMNLP 2023, pages 5274–5291, Singapore. Association for Computational Linguistics.