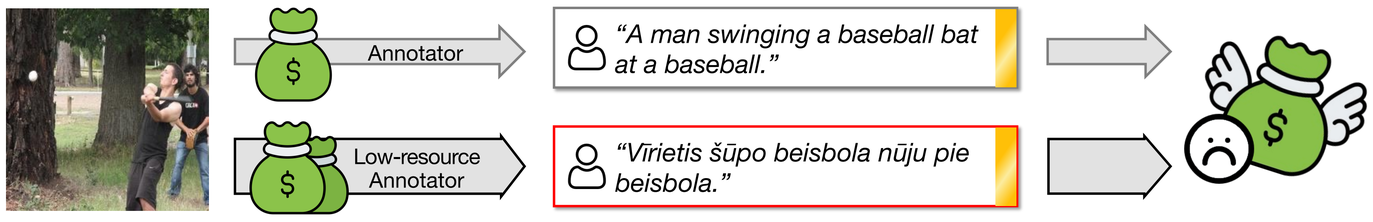
GPTs Are Multilingual Annotators for Sequence Generation Tasks

Juhwan Choi gold5230@cau.ac.kr
Eunju Lee dmswn5829@cau.ac.kr
Kyohoon Jin fhzh123@cau.ac.kr
YoungBin Kim ybkim85@cau.ac.kr

Chung-Ang University

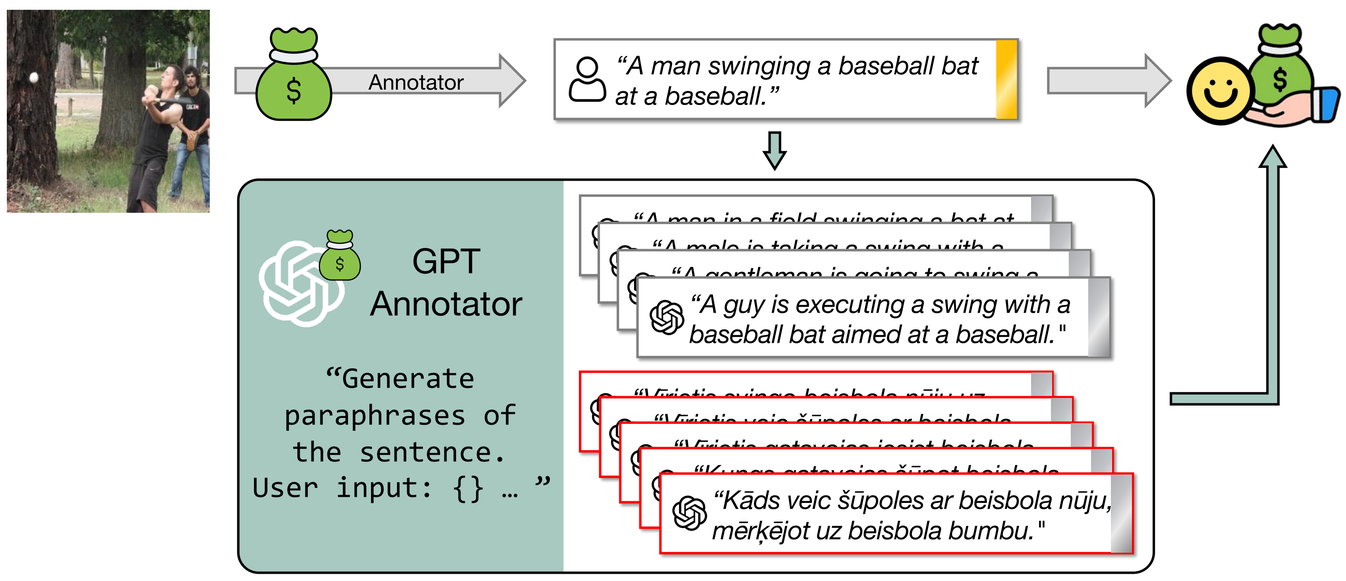
Motivation



- Data annotation is always hard and expensive
- The problem becomes more severe when it comes to low-resource languages

➔ We extend LLM-based annotation to **multilingual data annotation!**

Method



- Task: Image captioning - where 5 caption per each image is required
- We generate a set of paraphrases from **single** human-annotated data and translate them into designated language
- Our prompt design offers various benefits: debiasing, style of generated sentence, etc.

Dataset Construction

- We built image captioning dataset for **three low-resource languages**: Latvian, Estonian, Finnish
- Our analysis shows our method is superior compared to translation model regarding the quality of generated data

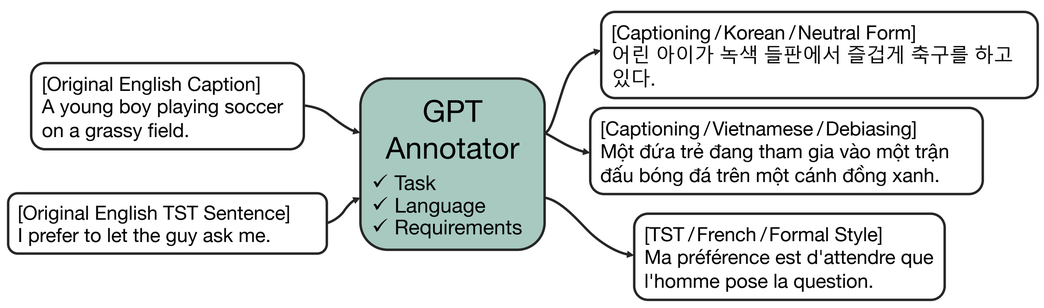
[Image]

***English Reference:**
There are four people playing tennis in doubles.

***NLLB (Machine-Translated):**
Divās grupās spēlē četri cilvēki.
(Four people play in two groups.)

***GPT Annotator w/ GPT-4:**
Četri cilvēki spēlē tenisu dubultspēlēs.
(Four people play tennis in doubles.)

[Captions]



Experiment

- We found our method is **cost-efficient, multilingual** data annotation method

Latvian	BLEU	ROUGE	METEOR	BERTS.	BARTS.
NLLB (Machine-Translated)	6.39	17.53	10.13	0.6803	-16.061
HRQ-VAE + NLLB	5.14	16.61	10.21	0.6728	-16.127
Google Translator	8.53	17.09	10.67	0.6848	-16.067
GPT Annotator w/ GPT-4	10.35	18.61	10.79	0.6911	-16.054
Estonian	BLEU	ROUGE	METEOR	BERTS.	BARTS.
NLLB (Machine-Translated)	4.97	13.12	7.89	0.6893	-15.409
HRQ-VAE + NLLB	3.37	7.84	5.87	0.6876	-15.409
Google Translator	6.04	12.51	8.75	0.7008	-15.408
GPT Annotator w/ GPT-4	6.62	13.47	9.22	0.7050	-15.407
Finnish	BLEU	ROUGE	METEOR	BERTS.	BARTS.
NLLB (Machine-Translated)	4.19	10.43	7.74	0.7122	-16.392
HRQ-VAE + NLLB	3.74	10.23	7.06	0.6965	-16.401
Google Translator	4.28	10.84	7.88	0.7128	-16.394
GPT Annotator w/ GPT-4	4.96	12.29	8.64	0.7143	-16.389