ChatGPT论文汇报

刘承伟

2023-05-17

ChatGPT论文汇报

▶ 文献阅读：

1. 《ChatGPT as a Text Simplification Tool to Remove Bias》University of York York, United Kingdom.

2. 《CODEIE: Large Code Generation Models are Better Few-Shot Information Extractors》Academy for Engineering & Technology, Fudan University

ChatGPT as a Text Simplification Tool to Remove Bias

- **Background** (Motivation)**:**
  - The presence of specific linguistic signals particular to a certain sub-group of people can be picked up by language models during training. This may lead to discrimination if the model has learnt to pick up on a certain group's language.
  - If the model begins to associate specific language with a distinct group, any decisions made based upon this language would hold a strong correlation to a decision based on their protected characteristic.

- **Solution**
  - Simplify language to one way of speaking to mitigation bias,

  while keeping the same meaning.

ChatGPT as a Text Simplification Tool to Remove Bias

▸ example

1. Automate resume scoring is becoming increasingly popular. This application could lead to unfairness if a person from a minority group was to have their resume declined due to the language they use.

2. Simplifying medical radiology reports into a form that is palatable to patients.

▸ Solution

1. ChatGPT (and GPT3.5 alone) excel at removing non-essential information and adding new information. Multilingual GPT models were also judged as well performing.

2. The above evaluation involves 18 benchmark datasets and 5 representative sentiment analysis tasks, and we compare ChatGPT with finetuned BERT and corresponding state-of-theart (SOTA) models on end-task.
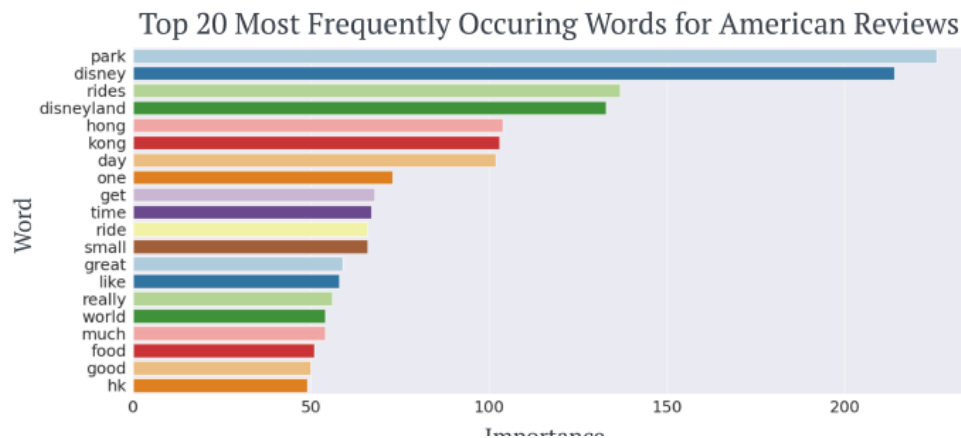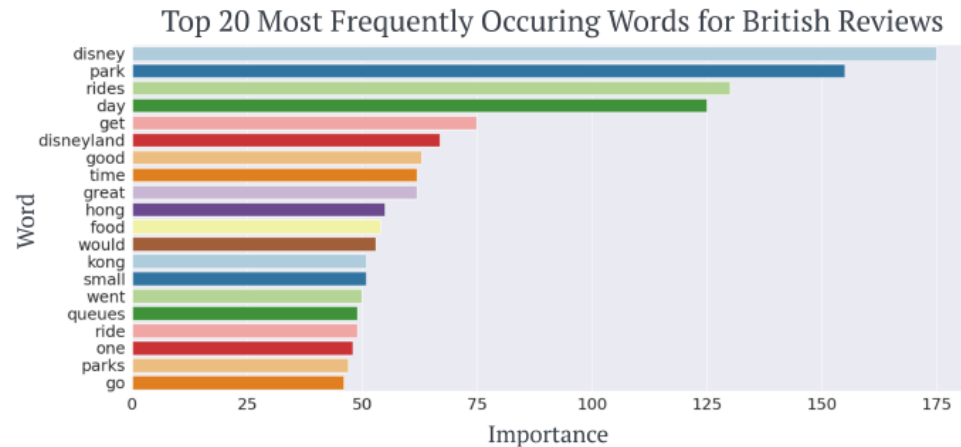
ChatGPT as a Text Simplification Tool to Remove Bias

Reviewer_Location	Review_Text	Disneyland_Branch
United States	Mickey and the Wounderous book was great and the kids love it. The new Ant...	Disneyland_HongKong
United Kingdom	We brought tickets before we left and got 2 days for less than the price of 1. After...	Disneyland_HongKong
. . .	. . .	. . .

ChatGPT as a Text Simplification Tool to Remove Bias

Top 20 Most Frequently Occuring Words for British Reviews

Top 20 Most Frequently Occuring Words for American Reviews

ChatGPT as a Text Simplification Tool to Remove Bias

Simplify "We really enjoyed Hong Kong Disney, however a few of the food outlets were closed. There was next to nothing for vegetarians or vegans. With so many of the food places not being open it made the queues worse. I think if your paying full price everything should be open. It was difficult to meet the characters, unlike when I went Florida and they were about everywhere. "

"We enjoyed Hong Kong Disney, but some food outlets were closed and there were limited options for vegetarians or vegans. The closure of many food places led to longer queues, which was disappointing given the full price of admission. Meeting characters was also difficult compared to Florida where they were more accessible."
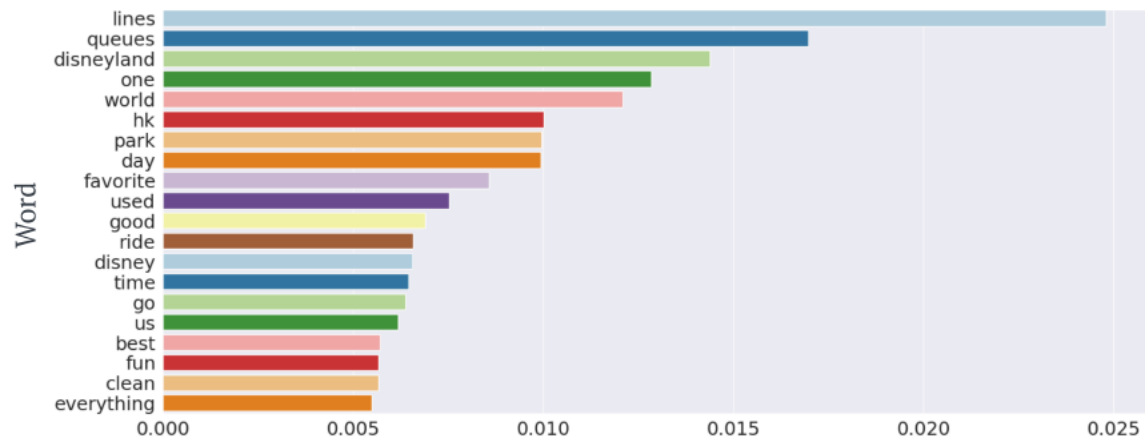
ChatGPT as a Text Simplification Tool to Remove Bias

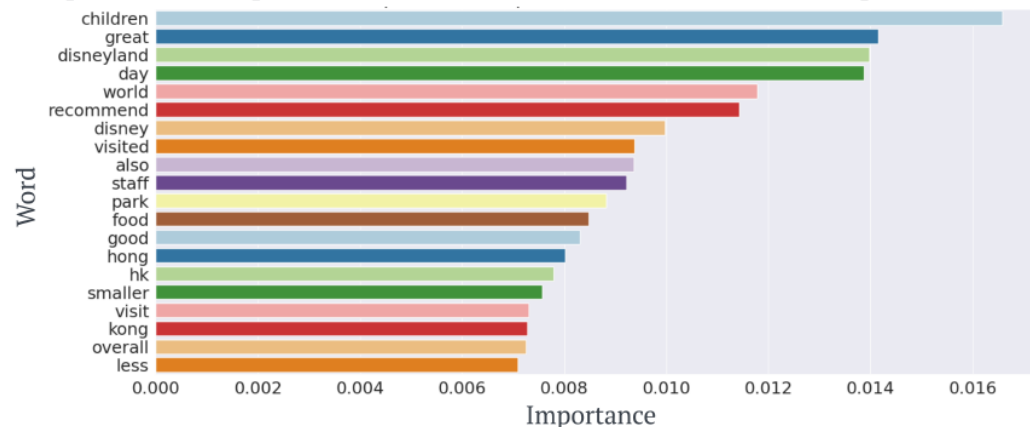| Dataset	Model	Accuracy	F1-score	Precision	Recall
Original	Naïve Bayes	0.80	0.79	0.91	0.71
	Random Forest	0.76	0.74	0.93	0.61
Simplified	Naïve Bayes	0.68	0.65	0.81	0.54
	Random Forest	0.59	0.58	0.66	0.51

ChatGPT as a Text Simplification Tool to Remove Bias

Top 20 Most Important Words for Classification of the Original Reviews

Top 20 Most Important Words for Classification of the Simplified Reviews

CODEIE: Large Code Generation Models are Better Few-Shot Information Extractors
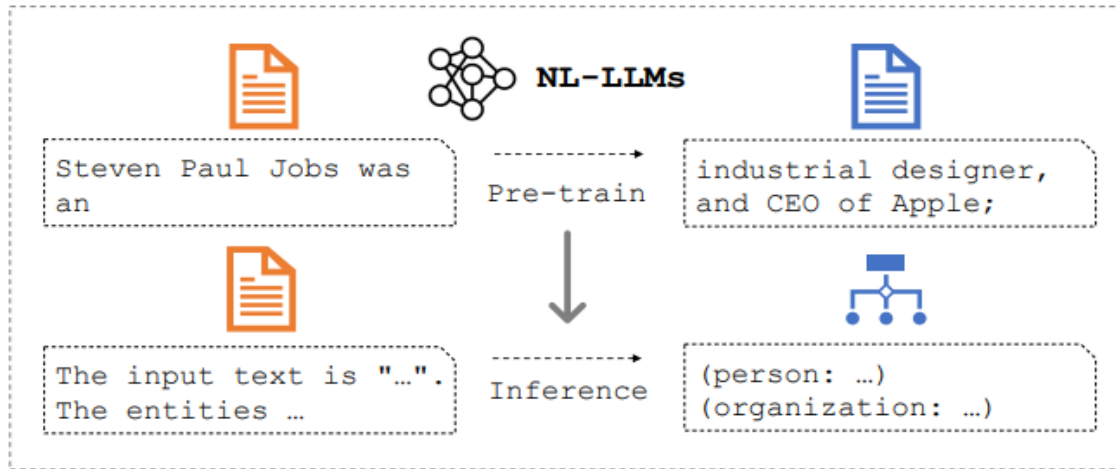
- **Background** (Motivation)：
  - It is non-trivial to perform information extraction (IE) tasks with NL-LLMs since the output of the IE task is usually **structured** and therefore is hard to be converted into **plain text.**
  - Existing linearizing approach still performs poorly under the few-shot scenario.
  - The abundant structured code information encoded in the pretrained Code–LLMs can benefit these IE tasks.

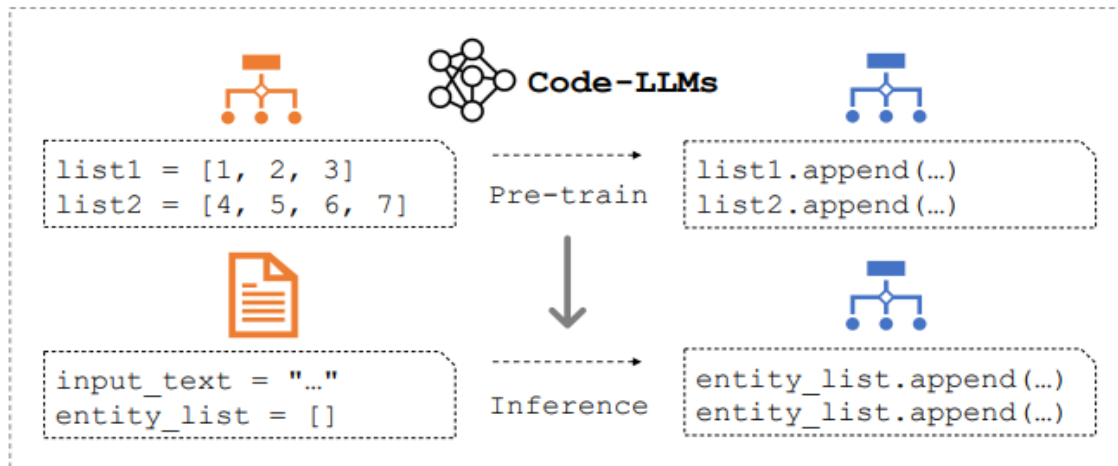- **Solution**
  - Code-LLMs model
  - Convert the **text–to–structure** IE task into a **structure–to–structure** code generation task.

CODEIE: Large Code Generation Models are Better Few-Shot Information Extractors

(a) Performing NER with NL-LLMs

CODEIE: Large Code Generation Models are Better Few-Shot Information Extractors

- **Solution**
    - Sample: "Steve became CEO of Apple in 1998 ." ->**Python**

```python
def named_entity_recognition(input_text):
    """ extract named entities from the input_text . """
    input_text = "Steve became CEO of Apple in 1998 ."
    entity_list = []
    # extracted named entities
    entity_list.append({"text": "Steve", "type": "person"})
    entity_list.append({"text": "Apple",\
                        "type": "organization"})
```

    - Feed the **code–style prompt** (the highlighted lines with light grey color) into **Code–LLMs** and get the structured prediction

CODEIE: Large Code Generation Models are Better Few-Shot Information Extractors

Model Type	Generative?	Extremely Large?	Structured Pre-train?	Few-Shot NER and RE Tasks
	Unified Framework	Few-shot Learning	Structured Task	
Pre. Models (e.g., UIE)	✓	✗	✓	✗
NL-LLMs (e.g., GPT-3)	✓	✓	✗	✗
Code-LLMs (e.g., Codex)	✓	✓	✓	✓

Table 1: A high-level comparison between previous IE Models, NL-LLMs and Code-LLMs. The bottom row illustrates our approach.

CODEIE: Large Code Generation Models are Better Few-Shot Information Extractors
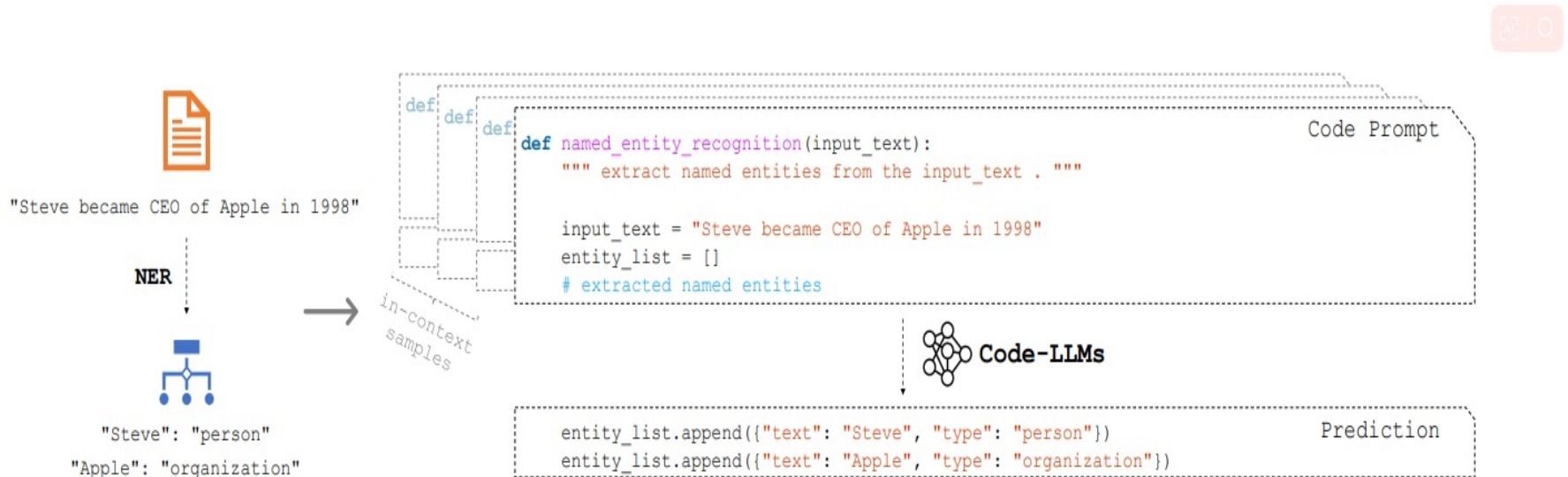
- **Findings：**
  - **Prompting Code-LLMs** with code-style inputs consistently outperforms fine-tuning pre-trained model for IE tasks, and prompting NL-LLMs (e.g., GPT-3) under few-shot settings.
  - With the **same LLM** (either NL-LLM or CodeLLM), the **code-style prompt** performs better than the **linearized text prompt**, demonstrating the advantage of representing structured targets with code.

  - With the **same prompt** (either natural language or code), the **Code–LLM** achieves better performance than the NLLLM (i.e., **GPT–3**), demonstrating the merits of performing IE tasks with Code–LLMs.

- **Solution(**Specific**)：**
  - Formulate the **NER** into **code generation task**

(a) Converting NER into code generation task

CODEIE: Large Code Generation Models are Better Few-Shot Information Extractors

- **Solution(**Specific**)：**
    - **Formulating IE Tasks into Code Generation Task** (Predict the subsequent code sequence given an incomplete piece of code.)
    - Reformulated into a list of dictionaries

CODEIE: Large Code Generation Models are Better Few-Shot Information Extractors

- **Solution(**Specific**)：**
  - It is nontrivial to perform IE tasks by prompting CodeLLMs without any **samples**, and it is necessary to let Code-LLMs be aware of a **few labeled samples** in typical few-shot settings.

CODEIE: Large Code Generation Models are Better Few-Shot Information Extractors

Struct text prompt

The text is "Steve became CEO of Apple in 1998 .". The named entities in the text: `((person: Steve)(organization: Apple))`

Natural text prompt

The text is "Steve became CEO of Apple in 1998 .". The named entities in the text: `"Steve" is "person". "Apple" is "organization".`

CODEIE: Large Code Generation Models are Better Few-Shot Information Extractors

Model	Prompt Type	Entity			
		CoNLL03	ACE04	ACE05-E	C
					Full Data
Pre. SoTA	-	**93.21**	86.84	84.74	
UIE-large	text	92.99	**86.89**	**85.78**	
					Few Shot
#shot (#sample)		5 (25)	2 (16)	2 (16)	
T5-base	text	$33.68_{\pm29.17}$	$7.25_{\pm12.00}$	$9.09_{\pm15.74}$	14
UIE-base	text	$70.37_{\pm0.54}$	$44.31_{\pm1.61}$	$39.71_{\pm0.91}$	4:
T5-large	text	$53.08_{\pm7.71}$	$24.67_{\pm5.26}$	$24.31_{\pm4.74}$	1(
UIE-large	text	$\mathbf{70.62}_{\pm3.22}$	$\mathbf{45.08}_{\pm3.63}$	$\mathbf{43.03}_{\pm2.26}$	4'
GPT-3	text	$68.84_{\pm1.29}$	$45.51_{\pm0.23}$	$48.93_{\pm0.49}$	3!
GPT-3	code	$81.00_{\pm1.49}$	$53.44_{\pm1.44}$	$52.98_{\pm1.53}$	5
Codex	text	$72.66_{\pm0.66}$	$49.58_{\pm1.37}$	$49.55_{\pm1.14}$	4'
Codex	code	$\mathbf{82.32}_{\pm0.37}$	$\mathbf{55.29}_{\pm0.37}$	$\mathbf{54.82}_{\pm2.09}$	5:

Thank you