SOOCHOW UNIVERSITY

工作汇报

刘承伟

2023-04-19

工作

- 文献阅读：

1. 《Zero-Shot Information Extraction via Chatting with ChatGPT》
   Beijing Jiaotong University; DAMO Academy, Alibaba Group, China
   - Transform the zero-shot IE task into a multi-turn question-answering problem with a two-stage framework (ChatIE).

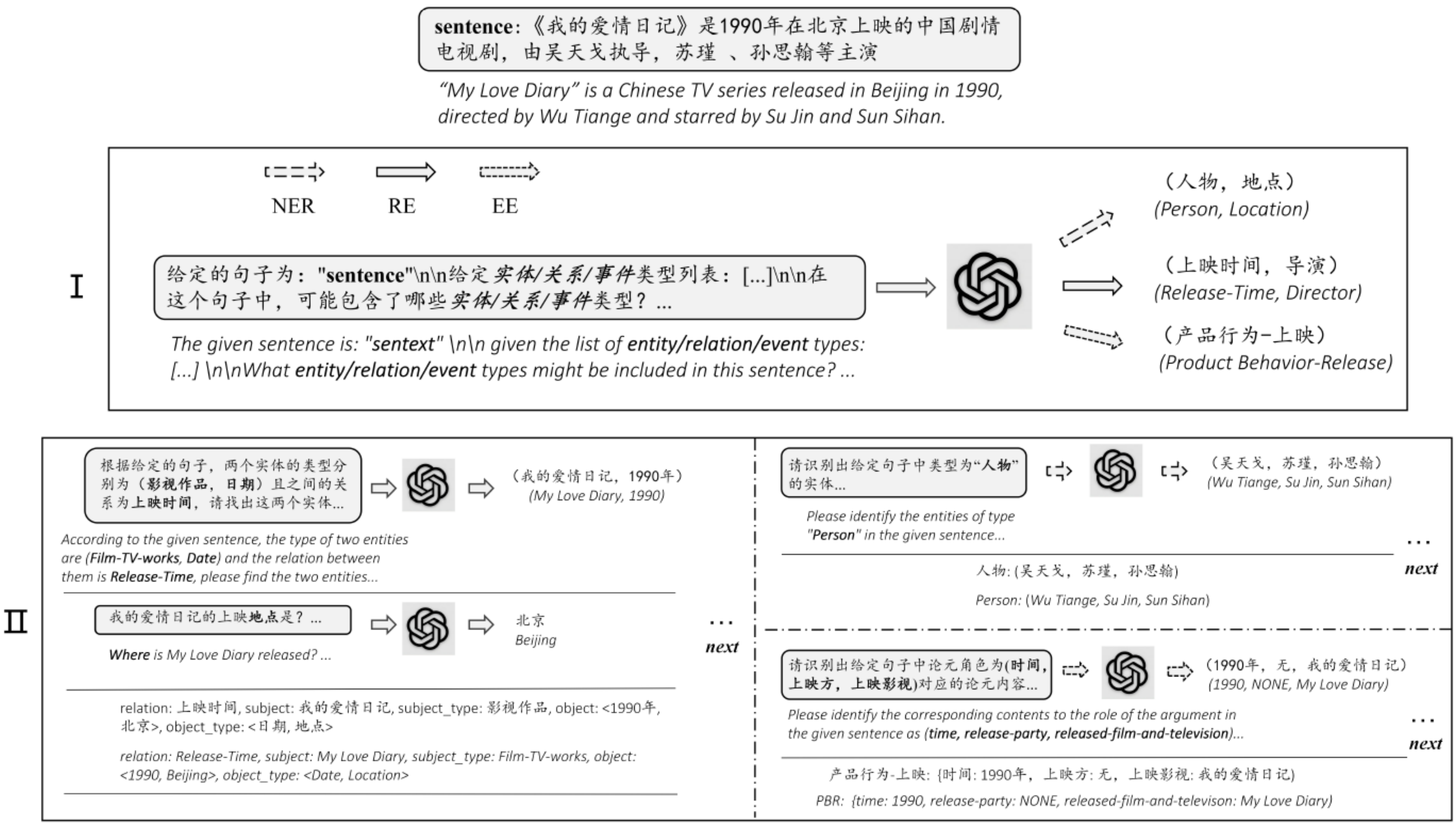
2. 《AugGPT:Leveraging ChatGPT for Text Data Augmentation》
   University of Georgia; South China University of Technology
   - Propose a text data augmentation approach based on ChatGPT (named AugGPT). AugGPT rephrases each sentence in the training samples into multiple conceptually similar but semantically different samples.

Zero-Shot Information Extraction via Chatting with ChatGPT

- **Motivation**
  1. Working with an enormous amount of labeling data is always hectic, labor-intensive, and time-consuming.
  2. Large-scale pre-trained language models (LLM), such as GPT-3, InstructGPT and ChatGPT, suggest that LLMs perform well in various downstream tasks even without tuning the parameters but only with a few examples as instructions.

- **Solution**
  1. In the first stage, aim to find out the corresponding element types (entities, relations,or events) that may exist in a sentence.
  2. In the second stage, perform a chained information extraction to each element type from Stage I.

Zero-Shot Information Extraction via Chatting with ChatGPT



Zero-Shot Information Extraction via Chatting with ChatGPT

- **Contributions**
  1. Each stage is implemented with a multi-turn QA process.
  2. In each turn, construct prompts based on designed templates and previously extracted information as input to ask ChatGPT.
  3. Finally, compose the results of each turn into structured data.
  4. Conduct extensive experiments on IE, NER, and EE tasks, including six datasets across two languages: English and Chinese. Empirical results show that two-stage framework instantiated on ChatGPT succeeds when the IE task is decomposed into multiple simpler and easier sub-tasks.

Zero-Shot Information Extraction via Chatting with ChatGPT

- **Specific implement**
  1. **Stage I**: Include only one turn of QA, utilize the task-specific TypeQuesTemplates and the list of element types to construct the question. Then combine the question and sentence as input to ChatGPT. If the sentence does not contain any element types, the system will generate a response with NONE Token.
  2. **Stage II**: Include multiple QA turns, design a series of specific ChainExtractionTemplates: define a chain of question templates (the extraction of an element may depend on another previous element) and the length of the chain is usually one.

Zero-Shot Information Extraction via Chatting with ChatGPT

1	Vanilla Prompt	Chat-based Prompt
STAGE I	***Question:*** I'm going to give you a sentence and ask you to identify the entities and label the entity category. There will only be 4 types of entities: ['LOC', 'MISC', 'ORG', 'PER']. Please present your results in list form. "Japan then laid siege to the Syrian penalty area and had a goal disallowed for offside in the 16th minute." Make the list like: ['entity name1', 'entity type1'],['entity name2', 'entity type2']....... - - - - - ***Expected Output:*** ["Japan", "LOC"], ["Syrian", "MISC"] ***Output:*** []	***Question:*** Given sentence: "Japan then laid siege to the Syrian penalty area and had a goal disallowed for offside in the 16th minute." The known entity types are: ['LOC', 'MISC', 'ORG', 'PER']. Please answer: What types of entities are included in this sentence? - - - - - ***Expected Output:*** LOC, MISC ***Output:*** LOC, MISC
STAGE II	None	***Question:*** According to the sentence above, please output the entities of 'LOC' in the form of list like: ['entity name1', 'entity type1'], ['entity name2', 'entity type2']...... - - - - - According to the sentence above, please output the entities of 'MISC' in the form of list like: ['entity name1', 'entity type1'], ['entity name2', 'entity type2']...... - - - - - ***Expected Output:*** ["Japan", "LOC"], ["Syrian", "MISC"]***Output:*** ["Japan", "LOC"], ["Syrian", "LOC"]

Table 3: Illustration of vanilla prompts vs our Chat-based prompts in terms of NER. The text highlighted with red represents the prompt template. The text following ***Question:*** represents the prompt that is used in ChatIE.

Zero-Shot Information Extraction via Chatting with ChatGPT

	RE						NER						EE					
	DuIE2.0			NYT11-HRL			MSRA			collnpp			DuEE1.0			ACE05		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1	P	R	F1	P	R	F1
fs-1	0.0	0.0	0.0	0.0	0.0	0.0	14.7	7.9	9.7	2.71	17.2	4.66	0.4	0.2	0.3	0.0	0.0	0.0
fs-5	0.0	0.0	0.0	0.0	0.0	0.0	34.5	10.3	15.5	2.53	16.65	4.38	0.2	0.6	0.3	0.0	0.0	0.0
fs-10	16.5	0.1	0.2	0.0	0.0	0.0	60.0	30.9	40.6	2.49	18.54	4.38	2.1	0.7	1.0	0.0	0.0	0.0
fs-20	41.4	0.4	0.8	3.4	2.7	0.5	63.4	44.8	52.5	2.48	19.36	4.41	1.7	0.8	1.1	4.6	0.1	0.2
fs-50	45.7	2.5	4.7	11.7	1.9	3.3	71.6	62.4	66.6	41.94	11.55	8.93	3.2	8.5	4.6	6.7	1.6	2.6
fs-100	50.8	7.2	12.0	34.8	6.2	10.6	81.3	76.1	78.6	50.26	24.97	32.89	8.7	12.0	10.1	8.0	4.9	6.0
full-shot	68.9	72.2	70.5	47.9	55.1	51.3	96.33	95.63	95.98	94.18	94.61	94.39	50.9	42.8	46.5	45.3	54.3	49.4
FCM	-	-	-	43.2	29.4	35.0	-	-	-	-	-	-	-	-	-	-	-	-
MultiR	-	-	-	32.8	30.6	31.7	-	-	-	-	-	-	-	-	-	-	-	-
single	17.8	7.7	10.7	10.8	5.7	7.4	56.3	**57.3**	56.8	61.4	43.0	50.6	61.7	77.5	68.7	18.2	23.9	20.7
ChatIE	**74.6**	**67.5**	**70.9**	**30.6**	**48.4**	**37.5**	**58.4**	57.0	**57.7**	**62.3**	**55.0**	**58.4**	**66.5**	**78.5**	**72.0**	**25.3**	**35.5**	**29.5**

Table 1: F1 score on six datasets over two languages.

AugGPT: Leveraging ChatGPT for Text Data Augmentation

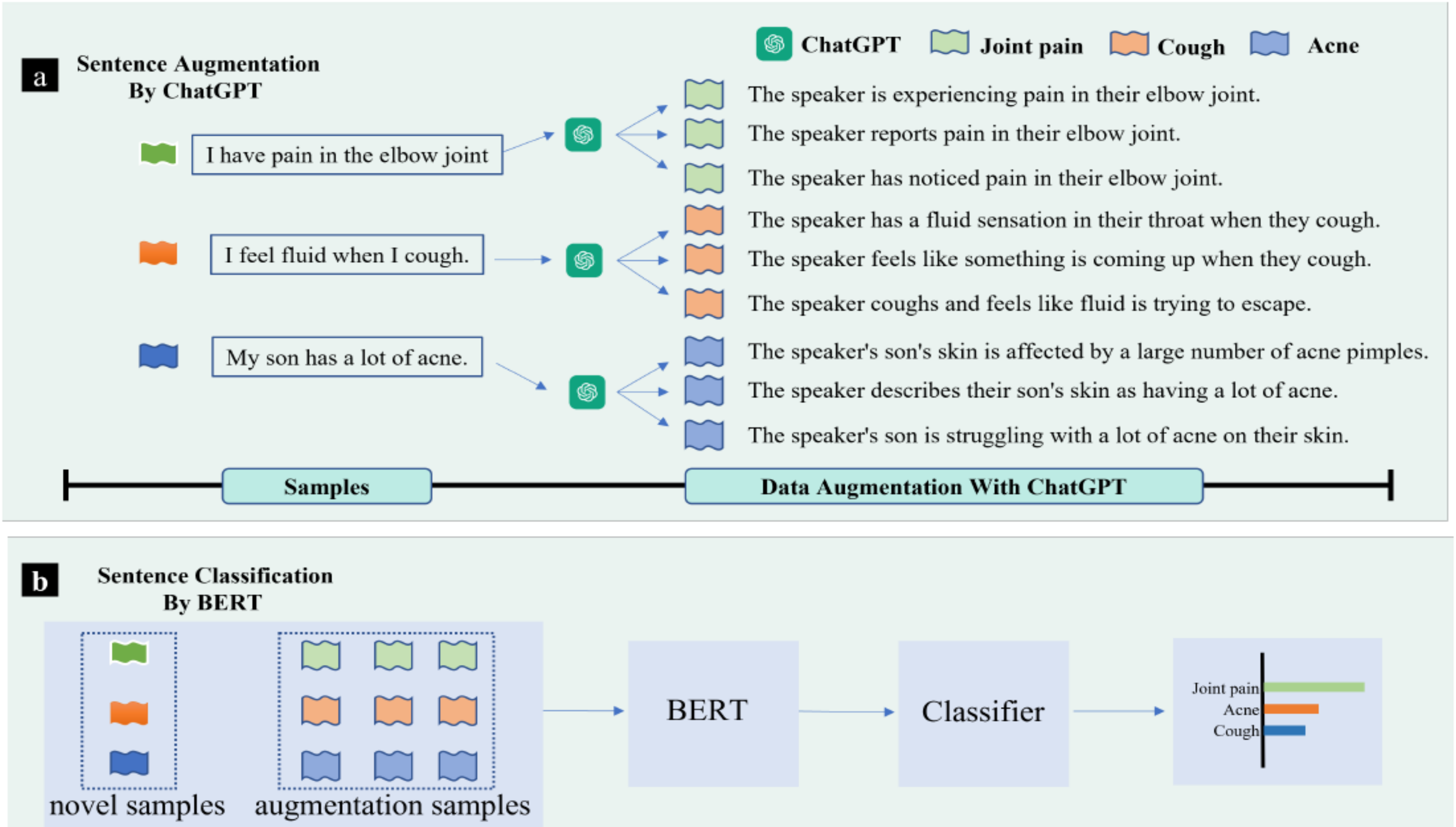
Motivation

1. The challenge of limited sample sizes in the few-shot learning scenario, where the data the target domain is generally much scarcer and of lowered quality
2. Current text data augmentation methods either can't ensure the correct labeling of the generated data (lacking faithfulness) or can't ensure sufficient diversity in the generated data (lacking compactness), or both.

Solution

1. Propose a new data augmentation method named AugGPT, which leverages ChatGPT to generate auxiliary samples for few-shot text classification.
2. Further investigation into the faithfulness and compactness of the generated text samples reveals that AugGPT can generate more diversified augmented samples while simultaneously maintaining their accuracy (i.e., semantic similarity to the original labels).

AugGPT: Leveraging ChatGPT for Text Data Augmentation



Zero-Shot Information Extraction via Chatting with ChatGPT

Original Text	Altered Text
The follow-up rate after 5 years was 85%.	1. The study showed a follow-up rate of 85% after 5 years. 2. The 5-year follow-up rate was found to be significantly high at 85%. 3. The participants had an impressive follow-up rate of 85% after 5 years. 4. The data indicated a follow-up rate of 85% 5 years post-study. 5. The 5-year follow-up rate was recorded as 85%, according to the findings. 6. The results of the study showed a follow-up rate of 85% after a 5-year period.
This means that the maturation of calluses was 27% faster in the LIPUS group.	1. The LIPUS group showed a 27% acceleration in callus maturation. 2. The results revealed that the callus maturation process was 27% faster in the LIPUS group. 3. The study found that the callus maturation rate in the LIPUS group was 27% quicker. 4. Callus maturation was 27% more rapid in the LIPUS group compared to others. 5. The callus maturation process in the LIPUS group was 27% more efficient, according to the results.
Test Registration: IRCT.ir IRCT2012071010230N1.	1. The trial was registered with the IRCT under the identifier IRCT2012071010230N1. 2. The study was officially registered with IRCT.ir, with the registration number IRCT2012071010230N1. 3. The trial was documented with IRCT.ir with the identification code IRCT2012071010230N1. 4. IRCT.ir served as the official registrar for the trial, with the registration number IRCT2012071010230N1. 5. The study was recorded with IRCT under the registration number IRCT2012071010230N1. 6. IRCT.ir recorded the trial with the identifier IRCT2012071010230N1.
While behavioral and technological interventions can lead to some modest improvements in glycemic control, these interventions have not performed much better than conventional prevention in achieving glycemic control.	1. The study found that although behavioral and technological interventions led to some slight improvements in glycemic control, they were not significantly more effective than typical care. 2. Despite the modest improvement in glycemic control through behavioral and technological interventions, they did not perform better than the standard care. 3. The results showed that while behavioral and technological interventions resulted in some minimal gains in glycemic control, they did not surpass the usual care in achieving glycemic control. 4. Although behavioral and technological interventions showed some improvement in glycemic control, they were not found to be significantly superior to the usual care. 5. The study showed that the usual care was not outperformed by behavioral and technological interventions in terms of achieving glycemic control, despite some small improvements.

Zero-Shot Information Extraction via Chatting with ChatGPT

- **Contributions**

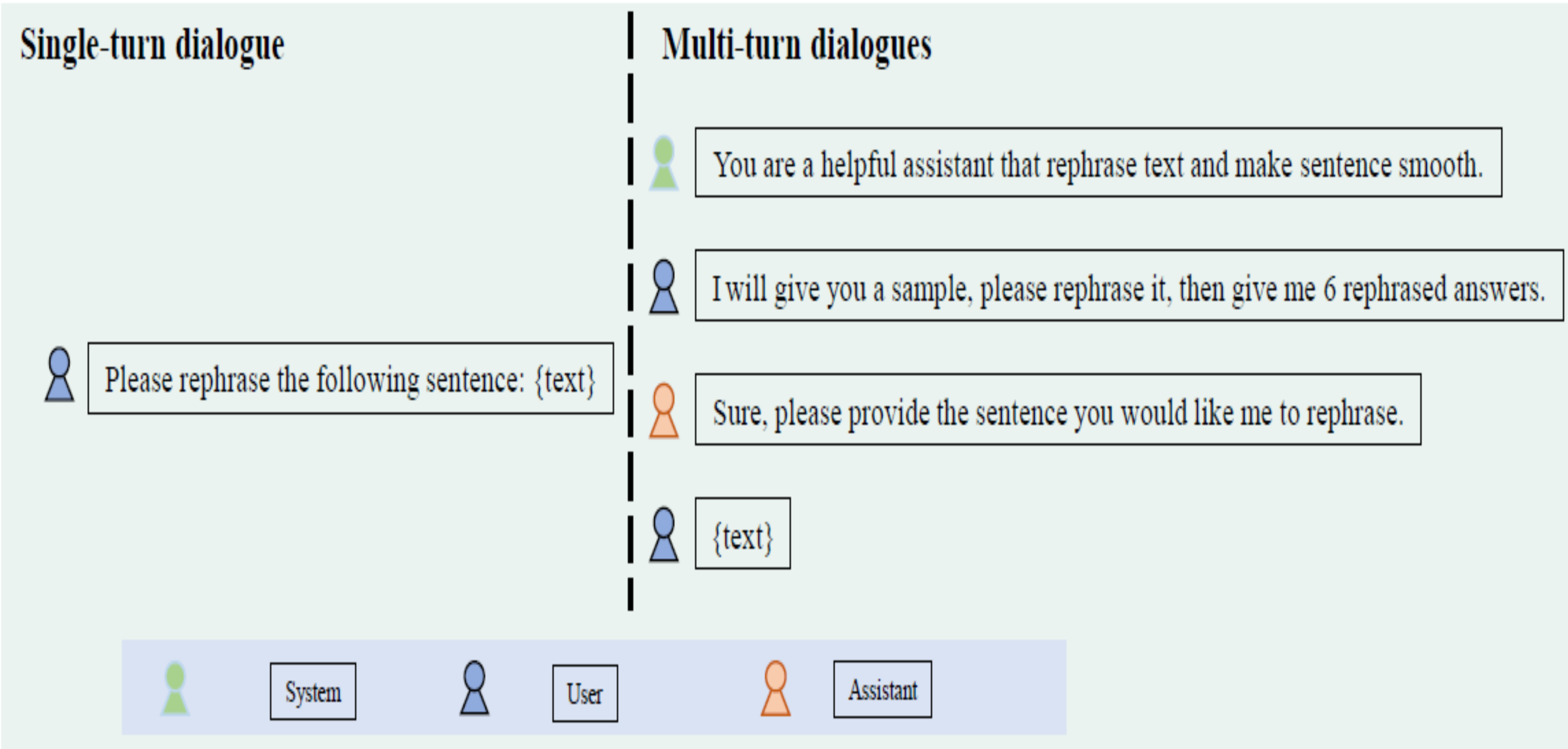
Algorithm 1 The framework of AugGPT for few-shot text classification.

Input: base dataset D_b and novel dataset D_n
Initialize: Initialized pre-trained BERT $model$
Definition: $D^{'}$ is the dataset with the base dataset D_b and augmented dataset D_n^{aug}, and *chatGPT_aug* is the data augmentation method based on ChatGPT
Parameters: Fine-tuning epochs of base dataset $epoch_b$, fine-tuning epochs of FSL $epoch_f$

```
for epoch in epoch_b do
    train(model, D_b)
end for
D_n^aug = chatGPT_aug(D_n)
for epoch in epoch_f do
    train(model, D_n^aug)
end for
```

Zero-Shot Information Extraction via Chatting with ChatGPT

- **Data Augmentation with ChatGPT**



Zero-Shot Information Extraction via Chatting with ChatGPT

Few-shot Text Classification

1. Objective function of few-shot learning consists of two parts: cross entropy and contrastive learning loss.
2. To make full use of the prior knowledge in the base dataset to guide the learning of the novel dataset, we introduce the contrastive loss function to make the sample representation of the same category more compact and the sample representation of different categories more separate.

Zero-Shot Information Extraction via Chatting with ChatGPT

Evaluation Metrics

1. **Embedding Similarity**: Adopt embedding similarity between the generated samples and the actual samples of the test dataset. Select cosine similarity to capture the distance relationship in the latent space. The cosine similarity measures the cosine value of the angle between two vectors.

2. **TransRate**: TransRate is a metric that quantifies transferability based on the mutual information between the features extracted by a pre-trained model and their labels, with a single pass through the target data. A higher TransRate could indicate better learnability of the data.

Zero-Shot Information Extraction via Chatting with ChatGPT

Data Augmentation and Ablation Study. The BERT + C indicates BERT with contrastive loss.

Data Augmentation	Amazon		Symptoms		PubMed20K	
	BERT	BERT + C	BERT	BERT + C	BERT	BERT + C
Raw	0.734	0.745	0.636	0.606	0.792	0.798
BackTranslationAug	0.757	0.748	0.778	0.747	0.812	0.83
ContextualWordAugUsingBert(Insert)	0.761	0.750	0.697	0.677	0.802	0.811
ContextualWordAugUsingBert(Substitute)	0.770	0.757	0.626	0.667	0.815	0.830
ContextualWordAugUsingDistilBERT(Insert)	0.759	0.762	0.707	0.747	0.796	0.796
ContextualWordAugUsingDistilBERT(Substitute)	0.787	0.766	0.667	0.646	0.797	0.800
ContextualWordAugUsingRoBERTA(Insert)	0.775	0.768	0.758	0.707	0.815	0.814
ContextualWordAugUsingRoBERTA(Substitute)	0.745	0.730	0.727	0.667	0.782	0.782
CounterFittedEmbeddingAug	0.754	0.741	0.667	0.626	0.805	0.805
InsertCharAugmentation	0.771	0.775	0.404	0.475	0.826	0.831
InsertWordByGoogleNewsEmbeddings	**0.816**	0.794	0.636	0.677	0.786	0.784
KeyboardAugmentation	0.764	0.766	0.545	0.505	0.809	0.815
OCRAugmentation	0.775	0.782	0.768	0.778	0.789	0.789
PPDBSynonymAug	0.691	0.690	0.697	0.758	0.795	0.829
SpellingAugmentation	0.727	0.736	0.697	0.707	0.808	0.811
SubstituteCharAugmentation	0.762	0.768	0.535	0.586	0.816	0.821
SubstituteWordByGoogleNewsEmbeddings	0.729	0.741	0.727	0.727	0.807	0.822
SwapCharAugmentation	0.762	0.766	0.475	0.485	0.797	0.801
SwapWordAug	0.771	0.766	0.687	0.727	0.798	0.794
WordNetSynonymAug	0.805	0.798	0.616	0.758	0.761	0.757
ChatGPT (2-shot)	0.753		0.980		0.748	
AugGPT	**0.816**	**0.826**	**0.889**	**0.899**	**0.835**	**0.835**

Thank you