** Contributions**

- **RECIPE4U (RECIPE for University) dataset**: Student-ChatGPT interaction dataset within the context of real-world English as a Foreign Language (EFL) writing education
- Baseline models for two subtasks with RECIPE4U: 1) intent detection 2) satisfaction estimation
- Investigation into the student’s interaction patterns: 1) students’ dialogue patterns 2) essay edit patterns

**RECIPE4U Dataset**

1) **Dialogue turn in RECIPE4U**
   - You are correct. The word "it" is unclear in this sentence. Can you think of a way to revise the sentence to make it clearer or more specific?
   - How about changing those two sentences to:
     - Even though tax is a burden, citizens constantly paying their tax.
   - Good job! That is a great improvement. Now let's look for any other unpleasant pronouns or repetitive words in the rest of your essay...

2) **Student's intent**
   - Request for Language Use
     - Request for Use
     - Request for Language Use
   - 4 out of 5

3) **Student's satisfaction**
   - Intent Detection
     - div1
     - div2
     - Example
     - Acknowledgement: AI: Please provide your essay. Student: [ESSAY]
     - Negotiation: The first two mistakes that you pointed out make no sense
     - Answer: AI: Can you please tell me what you learned in class? Student: Today I learned...
     - R for Translation: Can you translate it into Korean?
     - R for Confirmation: I bet there shouldnt be a lot of citation appear in the introduction, should it?
     - R for Language Use: Please point out the grammatical mistakes in the essay
     - R for Revision: Can you change this paragraph to be more effective to read?
     - R for Evaluation: Could you check if my essay has unity and coherence? Here is my essay.
     - R for Information: What are the common mistakes people do when writing an abstract?
     - R for Generation: Could you write it for me?
     - Question: Ahh… How many neurons do you have? I'm so curious about your structure and size.
     - Misc.
     - Statement: Also believe, think
     - Other: Hi there

4) **Essay edit history**
   - Original essay
     - ... First, the budget that will used was initially for all of us. We all citizens are paying tax, and they shout out many complaints such as 'It is too expensive.' Even though it is a burden, they pay the tax for mandatory. Using the tax for all citizen is right...
   - Edited essay
     - ... First, the budget that will be used was initially for all of us. We all citizens are paying tax, and they shout out many complaints such as ‘It is too expensive.’ Even though tax is a burden, citizens constantly paying their tax. Using the tax for all citizen is right...

**Experimental Results**

<table>
<thead>
<tr>
<th>Model</th>
<th>Intent Detection</th>
<th>Satisfaction Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>M-BERT</td>
<td>N/A</td>
<td>0.8344 (div1 3 cls)</td>
</tr>
<tr>
<td>XLM-R</td>
<td>N/A</td>
<td>0.7041 (div1 2 cls)</td>
</tr>
<tr>
<td>GPT-3.5</td>
<td>0.2 zero</td>
<td>0.1585 (div1 3 cls)</td>
</tr>
<tr>
<td></td>
<td>1.0 zero</td>
<td>0.1581 (div1 2 cls)</td>
</tr>
<tr>
<td></td>
<td>0.2 few</td>
<td>0.6202 (div1 3 cls)</td>
</tr>
<tr>
<td></td>
<td>1.0 few</td>
<td>0.5261 (div1 2 cls)</td>
</tr>
<tr>
<td>GPT-4</td>
<td>0.2 zero</td>
<td>0.7779 (div1 2 cls)</td>
</tr>
<tr>
<td></td>
<td>1.0 zero</td>
<td>0.7669 (div1 2 cls)</td>
</tr>
<tr>
<td></td>
<td>0.2 few</td>
<td>0.6991 (div1 2 cls)</td>
</tr>
<tr>
<td></td>
<td>1.0 few</td>
<td>0.6678 (div1 2 cls)</td>
</tr>
</tbody>
</table>

**Interaction Patterns**

**Human-like AI**

Hey ChatGPT you said...
Thank you ChatGPT

**Intelligent Peer**

that's what I selected too, but my professor marked that 'cannot' should also be selected

**Multilingual Entity**

Is it and B grammatical error?

번 문장에서 other 말고 다른 부분은 문법적 오류가 아니지 않아요?

(Am the rest of the sentence except for 'other' in sentence 1 free of grammatical error)

**Accept / Reject Feedback**

✔️ Student requested language use, revision, and information

❌ Feedback was trivial, untended, and hallucination

**Future Direction**

- Fine-tuned BERT achieve the highest results
- Ablation study on input conditions and essay masking to boost model performance
- Student & LLM with prompt recommendation
- Instructor & LLM with learning analytics

**Paper available at**

https://arxiv.org/abs/2302.00000

Jieun Han*, Haneul Yoo*, Junho Myung, Minsun Kim, Tak Yeon Lee, So-Yeon Ahn, Alice Oh
KAIST, South Korea
{jieun_han, haneul.yoo}@kaist.ac.kr