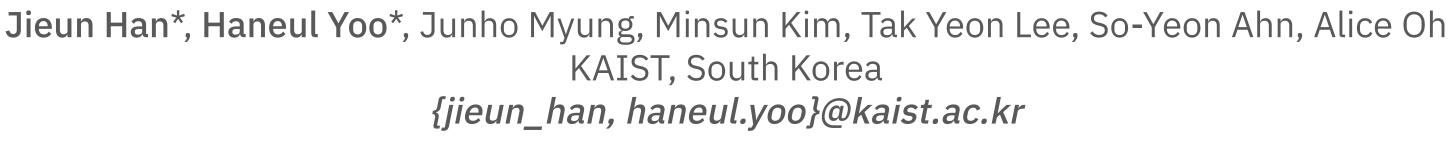


KAIST

Paper available at **Exploring Student-ChatGPT Dialogue in EFL Writing Education**



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	4) Essay edit history			
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?	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			
EFL student How about changing those two sentences to : Even though tax is a burden, citizens constantly paying their tax.	as 'It is too expensive.' Even though, it is burden, they pay the tax for mandatory. Using the tax for all citizen is right			



2) Student's intent

Request for Language Use

Good job! That is a great improvement. Now let's look for any other unclear pronouns or repetitive words in the rest of your essay...

4 out of 5

3) Student's satisfaction

... 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 tax is a burden, citizens **constantly paying their tax**. Using the tax for all citizen is right....

Intent		Evampla	Dist.	Avg. of	Satisfaction					
div1	div2	Example	(%)	Satisfaction	1	2	3	4	5	Total
Response	Acknowledgement	AI: Please provide your essay. Student: [ESSAY]	16.52	4.41	2	10	21	106	177	316
	Negotiation	The first two mistakes that you pointed out make no sense	5.07	3.89	4	8	18	32	35	97
	Answer	<i>AI: Can you please tell me what you learned in class?</i> <i>Student: Today I learned</i>	21.01	4.13	8	9	63	165	157	56
Request	R for Translation	Can you translate it into Korean?	1.25	3.71	2	2	4	9	7	24
	R for Confirmation	I bet there shouldn't be a lot of citation appear in the introduction, should it?	1.15	3.96	1	1	3	10	7	22
	R for Language Use	Please point out the grammatical mistakes in the essay	16.00	4.13	7	13	42	114	130	306
	R for Revision	Can you change this paragraph to be more effective to read?	8.57	4.05	6	6	28	58	66	164
	R for Evaluation	Could you check if my essay has unity and coherence? Here is my essay.	7.11	3.97	8	8	17	50	53	136
	R for Information	What are the common mistakes people do when writing an abstract?	13.12	4.17	6	10	28	97	110	251
	R for Generation	Could you write it for me?	3.14	4.07	1	7	5	21	26	60
	Question	Ahh How many neurons do you have? I'm so curious about your structure and size.	2.88	3.60	4	6	14	15	16	55
Misc.	Statement	Also believe, think	2.93	4.18	2	1	8	19	26	24
	Other	Hi there	1.25	3.83	4	0	4	4	12	402
						_				

Experimental Results

Model			Intent Detection		Satisfaction Estimation		
	temp.	shot	div1 (3 cls)	div2 (13 cls)	div1 (2 cls)	div2 (5 cls)	
M-BERT	Γ Ν/Α		0.8344	0.4291	0.9109	0.5794	
XLM-R	N/A		0.7041	0.2849	0.8591	0.4436	
GPT-3.5	0.2	zero	0.1585	0.2261	<u>0.8745</u>	0.4886	
	1.0	zero	0.1581	0.2251	0.8384	0.4604	
	0.2	few	0.6202	0.3053	0.7551	0.4435	
	1.0	few	0.5261	0.2047	0.7322	0.3968	
	0.2	zero	<u>0.7779</u>	<u>0.4891</u>	0.8626	<u>0.5639</u>	
GPT-4	1.0	zero	0.7669	0.4763	0.8582	0.5469	
	0.2	few	0.6991	0.4661	0.8188	0.4768	
	1.0	few	0.6678	0.4359	0.8069	0.4618	

- Fine-tuned BERT achieve the highest results
- Ablation study on input conditions and essay masking to boost model performance

Interaction Patterns

Human-like AI

Hey, ChatGPT you said.. Thank you, ChatGPT

Multilingual Entity Is 7 and 8 grammatical error?

7번 둔	장에서	other말고	! 다른
부분은	문법적	오류가	아니지
아 년		0}	?
(Isn't the	rest of the	sentence e	xcept for
'other' in	sentence 7	free of gra	mmatical
error)			

Intelligent Peer that's what i selected too, but my professor marked that 'cannot' should also have been selected

Accept / Reject Feedback

Student requested language use, revision, and information

X Feedback was **trivial**, unintended, and hallucination

Future Direction

- Student & LLM with prompt recommendation
- Instructor & LLM with learning analytics