



Week 3: Labelling Data

Alfonso Vieyra, Denise Wei,
Raha Ghazi, Tom Gao

Testing Labelling with Flan-T5-Large

	id	dialogue	summary	topic	sentiment
0	train_0	#Person1#: Hi, Mr. Smith. I'm Doctor Hawkins. ...	Mr. Smith's getting a check-up, and Doctor Haw...	get a check-up	positive
1	train_1	#Person1#: Hello Mrs. Parker, how have you bee...	Mrs Parker takes Ricky for his vaccines. Dr. P...	vaccines	positive
2	train_2	#Person1#: Excuse me, did you see a set of key...	#Person1#'s looking for a set of keys and asks...	find keys	positive
3	train_3	#Person1#: Why didn't you tell me you had a gi...	#Person1#'s angry because #Person2# didn't tel...	have a girlfriend	negative
4	train_4	#Person1#: Watsup, ladies! Y'll looking'fine t...	Malik invites Nikki to dance. Nikki agrees if ...	dance	positive
...
95	train_95	#Person1#: Let's plan, the dinner for the end ...	#Person1# and #Person2# are planning the class...	class dinner plan	positive
96	train_96	#Person1#: Hello Martin, long time no see. I h...	Martin tells #Person1# about his experience in...	daily casual talk	positive
97	train_97	#Person1#: Excuse me, do you have a few minute...	#Person1# surveys #Person2# about #Person2#'s ...	survey	positive
98	train_98	#Person1#: What kind of place shall we rent?\n...	#Person1# and #Person2# want a place near the ...	rent a place	positive
99	train_99	#Person1#: Hello! Can I help you?\n#Person2#: ...	#Person2# buys a bouquet of violet for #Person...	buy flowers	positive

Manual Analysis

- Very accurate over 100 manually analyzed rows
- Fairly slow to label - 100 rows took 5 minutes
 - Can the process be optimized?
- Majority of dialogue seems to be positive
 - About 80% of the 100 rows were labelled as positive

...

Next Steps

- How do we know the labelling is accurate enough?
- Are there other models we should try for labelling?
 - GPT-4 (cost), other models with Zero-Shot Learning capabilities
- Should we fine tune a model for this task?

...