

Utterance	Prediction Set	LLM output
how many more miles can i drive on my current tank	gas, mpg	I understand you are asking about your car's fuel. Are you asking about the number of miles you can drive on your current tank or the miles per gallon?
can i redeem my points for flights	redeem rewards, rewards balance	I understand you would like to redeem your points for flights. Are you looking to check your rewards balance or redeem your rewards?
are there any meetings between 2 and 5 pm	schedule meeting, meeting schedule	I understand you are looking for a schedule of meetings. Are you looking for a specific time range or a general schedule?
is the marathon on friday added to my calendar	calendar, calendar update	I understand you are referring to your calendar. Is it the marathon on friday that you are referring to or do you need an update on your calendar?
why didn't my card work	expiration date, card declined	I understand you are having issues with your card. Is it because of the expiration date or was it declined?

Table 7: Sample of LLM outputs on the C150 dataset.

D Appendix: Intent distributions

We here present the intent distributions for all public datasets across train, calibration and test splits.

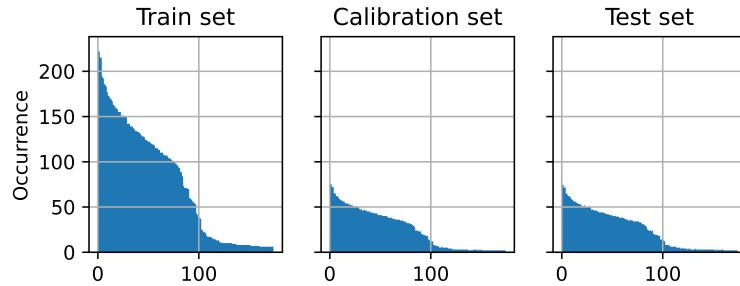


Figure 3: Intent distribution in ACID data set.

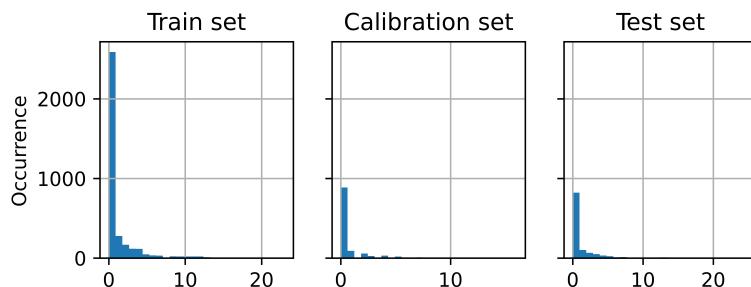


Figure 4: Intent distribution in ATIS data set.

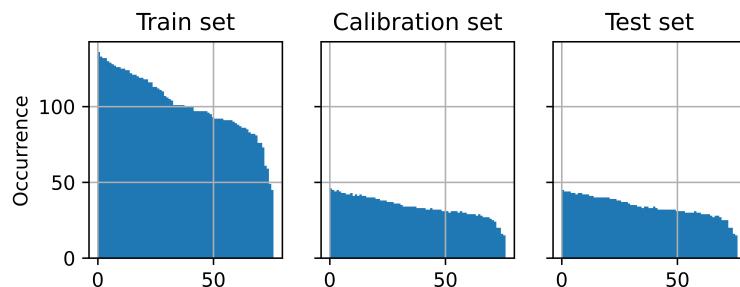


Figure 5: Intent distribution in B77 data set.

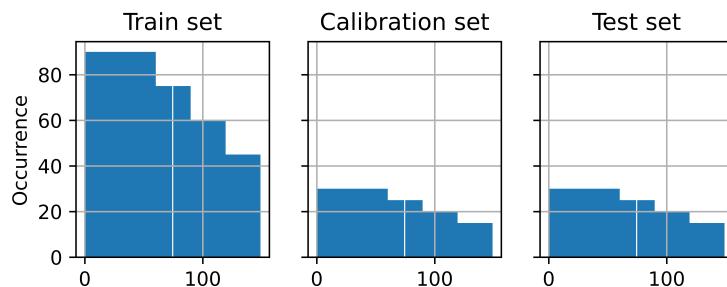


Figure 6: Intent distribution in C150-IS data set.

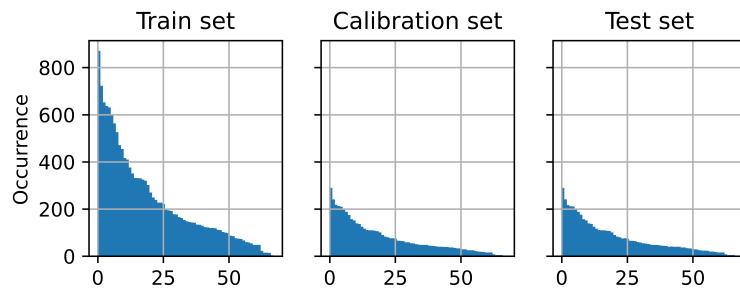


Figure 7: Intent distribution in HWU64 data set.

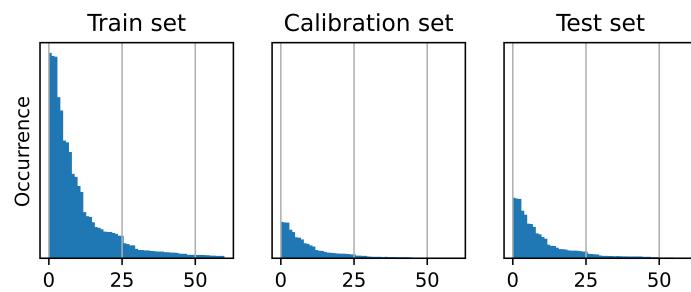


Figure 8: Intent distribution in IND data set.

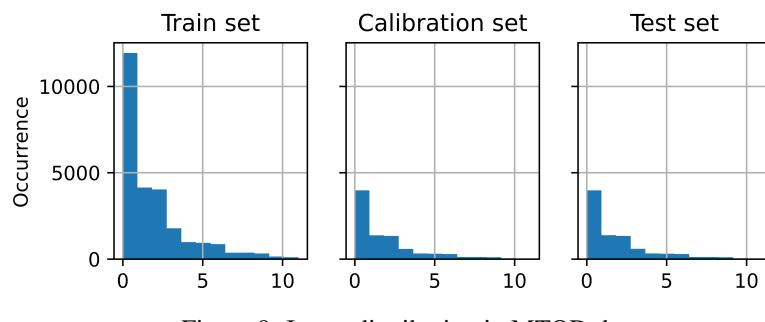


Figure 9: Intent distribution in MTOD data set.