

Chatbot Report

Team Members:

Madhav Adhikari
Sai Neeraj Chandragiri
Nirupavardhan Lingareddy

Dataset:

For training we have used few custom responses like “Hi”, “Hello”, “How are you?”. And, from a file “daily_dialog_data.csv” which is downloaded from the Hugging Face Datasets API.

Preprocess:

Due to an API(https://huggingface.co/datasets/daily_dialog?viewer_api=true) limitation allowing only 100 rows per request, the program makes three requests with different offset values (0, 100, and 200) to retrieve a total of 300 rows. Here, one row has one full conversation. For each request, processed the JSON response to extract the dialog data. Selected only the columns required for the chatbot (e.g., dialog columns). Cleaned the text by removing punctuation and any other unwanted characters. And, saved the cleaned data to a CSV file.

Custom Feature:

For custom feature, we used logic adapters like BestMatch, TimeLogicAdapter and MathematicalEvaluation.

- **BestMatch Logic Adapter:** Enables accurate matching of user inputs with predefined patterns for more precise responses, improving user interaction.
- **TimeLogicAdapter:** Handles time-related queries effectively, enhancing context awareness and enabling tasks like scheduling and reminders.
- **MathematicalEvaluation:** Performs mathematical calculations swiftly, offering users instant assistance with arithmetic operations and problem-solving.

POS Tags:

We can use SQLite Viewer or use any other sites on online to check conversion file of the SQLite Database.

id	text	search_text	conversation	created_at	in_response_to	search_in_response_to	pe
31	Thats true but then you h...	NNS:true TO:get DT:evide...	training	2024-04-01 21:24:19.592...	The radio has too many c...	DT:broadcasting RB:many...	
32	Are you all right	RB:change	training	2024-04-01 21:24:19.592...	Thats true but then you h...	NNS:true TO:get DT:evide...	
33	I will be all right soon I wa...	DT:right JJ:soon VBD:frig...	training	2024-04-01 21:24:19.592...	Are you all right	RB:change	
34	Dont worryHe is an acrobat	NNP:worryhe DT:athlete	training	2024-04-01 21:24:19.592...	I will be all right soon I wa...	DT:right JJ:soon VBD:frig...	
35	I see	PRP:perceive	training	2024-04-01 21:24:19.607...	Dont worryHe is an acrobat	NNP:worryhe DT:athlete	

In this conversation our chatbot used the POS tags “NNS”, “RB”, “DT”, “NNP”, and “PRP”.

- NNS: Noun, plural (*true* in this example)
- RB: Adverb (*Change* in this example)
- DT: Determiner (*right*, *athlete* in this example)
- NNP: Proper noun, singular (*worryhe* in this example)
- PRP: Personal pronoun (*perceive* in this example)