

TOPIC STEERING OPEN-ENDED CHATBOT

B. Tech. Project End Sem Report

-Submitted by

YASH BURBURE **112003022**

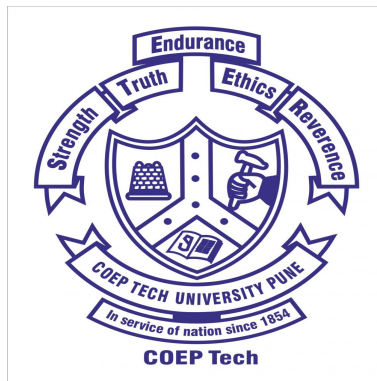
YASH CHANDAK **112003023**

JAYRAJ CHARANE **112003024**

Under the guidance of

Prof. ARCHANA PATIL

COEP Technological University, Pune



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
COEP TECHNOLOGICAL UNIVERSITY, PUNE-5

May 2024

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING,
COEP TECHNOLOGICAL UNIVERSITY, PUNE-5**

CERTIFICATE

Certified that this project titled, “TOPIC STEERING OPEN-ENDED
CHATBOT” has been successfully completed by

YASH BURBURE 112003022

YASH CHANDAK 112003023

JAYRAJ CHARANE 112003024

and is approved for the partial fulfillment of the requirements for the degree
of “B.Tech. Computer Engineering”.

SIGNATURE

Mrs. A. B. Patil

Project Guide

Department of CSE

COEP Tech Pune,

Shivajinagar, Pune - 5.

SIGNATURE

Mr. P. K. Deshmukh

Head

Department of CSE

COEP Tech Pune,

Shivajinagar, Pune - 5.

Abstract

Many real-world conversations in open-domain applications involve specific objectives such as providing recommendations, offering psychotherapy, or facilitating education. In this project, we aimed at developing a chatbot capable of steering conversation smoothly towards set topic. Our approach involves selecting a response topic such that the generated response steers the conversation progressively in the direction of the target topic. Evaluation results, both quantitative and human-based, demonstrate the effectiveness of the proposed approach.

Contents

List of Tables	4
List of Figures	5
1 Introduction	1
2 Literature Review	3
3 Research Gaps and Problem Statement	5
3.1 Research Gaps	5
3.2 Problem Statement	5
4 Proposed Methodology/ Solution	6
5 Experimental Setup	10
6 Results and Discussion	11
7 Conclusion	17

List of Tables

6.1	Fine-tuned DialoGPT across various metrics	13
-----	--	----

List of Figures

1.1	Example of a target-oriented conversation with natural flow. .	2
4.1	1st hop - 1st hop intersection.	7
4.2	1st hop - 2nd hop intersection.	7
4.3	1st hop of current topic.	8
6.1	Analysis of methods for finding next suitable topic.	12
6.2	Successful chat 1.	14
6.3	Successful chat 2.	15
6.4	Chat with an abrupt transition.	15
6.5	Fine-tuned DialoGPT fails to generate response containing the next keyword.	16

Chapter 1

Introduction

In recent years, chatbots have emerged as indispensable tools in numerous domains, offering invaluable support in customer service, personal assistance, and beyond. These conversational agents represent a bridge between users and information, leveraging natural language processing (NLP) techniques to engage in human-like interactions. Despite their widespread adoption, a significant challenge persists: the ability to seamlessly guide conversations towards predefined topics.

The aim of this project is to tackle this challenge head-on by developing an advanced chatbot capable of deftly steering conversations towards specific target topics, all without the user's explicit direction. Whether the target is a particular subject matter, a product recommendation, or even a service offering, the chatbot will autonomously guide the dialogue towards this goal.

By analyzing user inputs and conversation dynamics, the chatbot dynamically generates responses that subtly transition the discussion towards the target topic. Through a combination of context understanding and topic modeling, the chatbot adeptly navigates the conversation landscape, ensuring a smooth and natural progression towards the desired subject matter.

```
TARGET - movie

USER > i am bored.
BOT  > what do you like to do?

USER > i like to swim.
BOT  > i love swimming too. what else do you like?

USER > i love to sing and read.
BOT  > do you also like watching movies?
```

Figure 1.1: Example of a target-oriented conversation with natural flow.

This project report serves as a comprehensive exploration of our solution, documenting the development process, evaluation methodologies, and in-depth analysis of the results. Additionally, it delves into the underlying technologies and algorithms employed, shedding light on the intricacies of chatbot development and conversational AI.

Looking ahead, the report also discusses potential avenues for future research and enhancement. As the field of NLP continues to evolve, so too will our chatbot, with ongoing efforts focused on improving accuracy, scalability, and user experience. Through this project, we aim to contribute to the advancement of conversational AI and pave the way for more intelligent, intuitive interactions between humans and machines

Chapter 2

Literature Review

The past research on chatbots mainly focused on open-domain and task oriented chat systems. Task-oriented systems focus on achieving specific goals, such as providing information or facilitating negotiations. In contrast, chat-oriented systems aim to model open-domain conversations without predefined objectives. There has been comparatively very less research which aims to implement goal specific dialogue in open-ended chatbots.

In the paper by Tang et al. (2019) [7], the authors proposed an approach to target-guided open-domain conversation. By breaking down the system into modular components and introducing coarse-grained utterance keywords, they achieved precise control over the intended content of each response. They used a discourse-level rule that encourages these keywords to approach the end target during the conversation and employ turn-level supervised learning for smooth transitions between dialogue turns. They proposed a Kernel method which used a RNN model to predict the next keyword based on the conversation history. For producing responses they employed keyword-augmented response retrieval.

In their follow-up study, Qin et al.(2020) [5] built upon the work of Tang et al. (2019) by introducing enhancements to the proposed solution. They introduced a dynamic knowledge routing network (DKRN) which consider

the semantic knowledge relations among candidate keywords and the conversation context for keyword prediction. Utilizing the relation graph between candidate keywords, DKRN predicts the next keyword, effectively reducing search space and enhancing prediction accuracy. They too used a keyword-augmented response retrieval module for producing responses. Through supervised learning on chat data, DKRN demonstrates accurate keyword prediction, facilitating smooth transitions and effective target guidance. The comparative evaluation showed a better success rate in achieving the target topic with more smoothness.

In a comparative human evaluation, DKRN had a target achievement ratio of 88%, whereas the Kernel model reached 58% with conversation limit set to 8 turns. DKRN also received a smoothness score of 3.22, while the Kernel model scored 2.48.

Further study by Kishinami et al. (2022) [1] proposed an approach where the authors used ConceptNet to get the candidate keywords for next response based on the target keyword. The candidate keyword having highest similarity score with current keyword is selected as the next keyword for response. They used fine tuned BlenderBot(3B) to generate responses given the conversation history and keyword. The comparative evaluation showed better smoothness in guiding the conversation compared to earlier methods. However, the target topic achievement ratio was lower than previous methods.

Chapter 3

Research Gaps and Problem Statement

3.1 Research Gaps

The idea of directing open-domain conversations towards particular topics has been relatively underexplored. Some approaches excel in facilitating smooth transitions but struggle to reach the intended topic while some are good at achieving the target topic but may lack a natural flow in dialogue.

3.2 Problem Statement

The problem statement revolves around the exploration and advancement of open-ended chatbot systems with the capacity to guide conversations towards predefined topics smoothly. This gap inhibits the effectiveness of chatbots in scenarios where subtly steering conversations towards conversation specific goals, such as in advertisements, surveys, etc.

Chapter 4

Proposed Methodology/ Solution

Our approach uses ConceptNet [6] to find the topic for the response and DialoGPT to generate the response based on the conversation history and the predicted topic.

The proposed solution consists of the following steps:

1. **Finding User Utterance Keyword:** We employ POS tagging to identify candidate keywords from the utterance. Among these candidates, we select the keyword most closely associated (using relatedness scores) with the target keyword. We assign weights of 2 for nouns, 1 for verbs, 0.5 for adjectives, and 0 for other parts of speech. If no keyword is detected, we fallback to using the keyword from the previous utterance.
2. **Finding Candidate Topics for Response:**
 - (a) Here first we try to find the common words from words directly linked with current topic and directly linked with target topic in ConceptNet (i.e. 1st hop - 1st hop intersection).

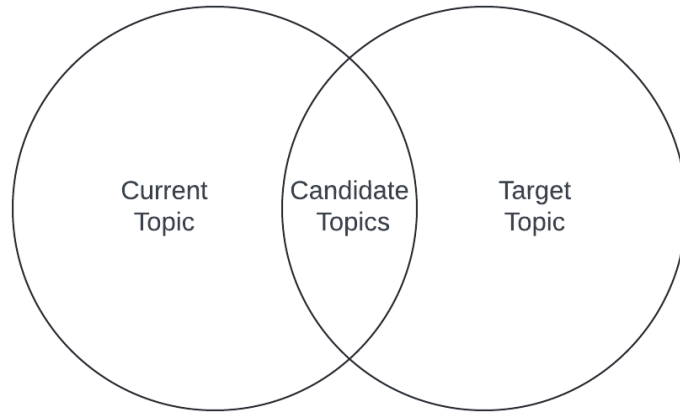


Figure 4.1: 1st hop - 1st hop intersection.

- (b) If 1st hop - 1st hop intersection is empty then we find the common words between first hop of current topic and second hop of target topic in ConceptNet (1st hop - 2nd hop intersection).

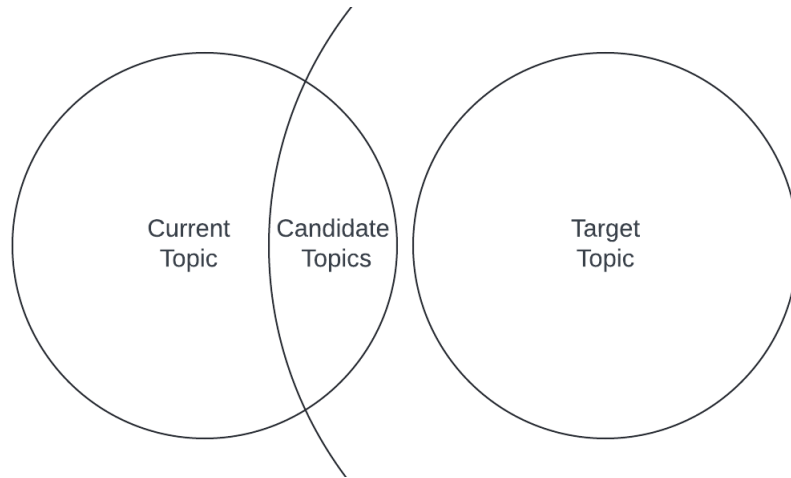


Figure 4.2: 1st hop - 2nd hop intersection.

- (c) If 1st hop - 2nd hop intersection is also empty then we find all the words directly linked with current topic in ConceptNet (Complete 1st hop of current topic).

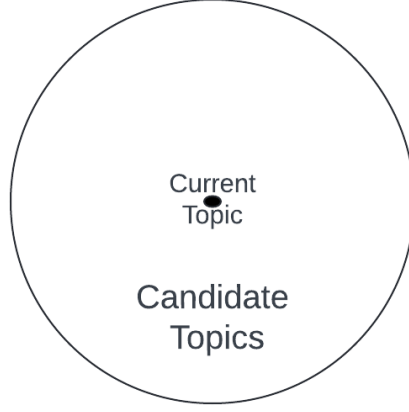


Figure 4.3: 1st hop of current topic.

3. Next Keyword Selection : We use relatedness value as a metric to measure distance between two topics. We check the relatedness value provided by ConceptNet for all the candidate topic with the target topic. We select the word with the highest relatedness value as the next keyword.

We conducted an analysis where we got better intermediate words using relatedness value than using cosine similarity.

4. Response Generation: We utilize the fine-tuned DialoGPT to generate responses, considering both the conversation history and the keyword provided for crafting the response.

For Fine-tuning DialoGPT,

- First we find keyword (utterance topic) for all utterances the dataset. For this, we utilize TF-IDF (Term Frequency-Inverse Document Frequency) and POS (Part of Speech) tagging methods to find the most important word from the utterance. We give POS weights as follows, 2 for noun, 1 for verb, 0.5 for adjective and 0 for others. Finally, we lemmatize the extracted keywords.

- Now, we fine-tune the DialoGPT-medium (345M) model on the dataset with input as (conversation history, next utterance keyword) and output as next utterance. Conversation history includes the past 5 utterances of conversation.

By following this methodology, we aim to develop a robust Topic Steering Open-ended Chatbot capable of steering conversations towards predefined targets effectively and generating coherent and engaging responses.

Chapter 5

Experimental Setup

We utilized Kaggle’s P100 GPU, equipped with 16GB of dedicated graphics memory, alongside 29GB of RAM for training and experimentation. Our cloud storage platform was Kaggle’s storage service. For our deep learning environment, we relied on the PyTorch framework along with the Transformers library to facilitate working with the pre-trained DialoGPT model.

The dataset was derived from the PersonaChat dataset [8] where crowdworkers engaged in natural conversations while embodying assigned personas. The conversations cover a broad range of topics such as work, family, personal interest, etc. and the discussion topics also change frequently. This made it suitable for training models to generate natural conversational flows. The length of the training, validation, and testing dataset was 109697, 12250, and 552 respectively. We experimented with different batch sizes and found that the most optimal results were achieved with a batch size of 4. Training was executed for one epoch, employing the AdamW optimizer with a learning rate set to $5e-5$.

Chapter 6

Results and Discussion

Method used for finding the next topic for the response

We did analysis of which method should be used to select the next topic for response. we selected random 50 start topic and random 50 target topic and found the next keyword using 4 ways,

1. only 1st hop method and cosine sim,
2. 1st hop method and relatedness value,
3. intersection method and cosine sim,
4. intersection method and relatedness value

Manual evaluation was done to see satisfactory and unsatisfactory intermediate topic words. Intersection method with relatedness value method gave the most number of satisfactory intermediate topics.

	A	B	C	D	E	F	G
1	word1	word2	og_cossim	og_relatedness	intersection	cossim	relatedness
2	book	movie	book	blockbuster	18 (1-1)	musical	script
3	school	dance	dance	dance	14 (1-1)	dance	dance
4	dance	basketball	dancer	athletic	3 (1-1)	sport	sport
5	pet	pedigree	doga	genetics	2 (1-1)	dog	dog
6	egg	farm	farms	farms	22 (1-1)	farms	farms
7	music	kathak	piphat	jhankaradhvani	2 (1-1)	lahara	dance
8	piano	dance	harp	music	5 (1-1)	musical	music
9	chocolate	dance	dessert	shake	5 (1-1)	bars	shake
10	bank	space	guard	storage	18 (1-1)	office	area
11	burger	computer	food	dabeli	8 (1-2)	food	patty
12	computer	sand	rice	slab	3 (1-1)	window	power
13	detective	drama	tv	thriller	2 (1-1)	howcatchem	whodunit
14	astronomy	badminton	astronomy	eunomia	24 (1-2)	fish	rayleigh
15	perfume	india	lavender	aryballos	2 (1-1)	mixture	sandalwood
16	swimming	movie	lake	swimmingness	2 (1-1)	moving	moving
17	panda	youtube	uk	pandamonium	2 (1-2)	uk	uk
18	violin	photographer	musical	violinist	3 (1-2)	musical	musical
19	military	ganga	sangar	gharry	9 (1-2)	standard	sainik
20	rabbit	hockey	football	football	4 (1-1)	british	game
21	dragon	china	phantom	mahjong	25 (1-2)	constellation	mahjong
22	nasa	china	australia	australia	1 (1-1)	australia	australia
23	astronaut	snake	person	tang	8 (1-2)	person	tang
24	diabetes	marathon	diabetes	meglitinide	1 (1-2)	hormone	harmone
25	pumpkin	mexico	australia	pepita	7 (1-2)	australia	us
26	cotton	maharashtra	kalamkari	surat	10 (1-2)	kalamkari	surat
27	forest	football	usa	rangership	4 (1-1)	field	field
28	school	computer	desk	desk	4 (1-1)	desk	desk
29	amazon	panda	parrot	fang	2 (1-2)	woman	river
30	cancer	vegan	cancerian	animal	1 (1-1)	person	person
31	money	cancer	charity	fundraising	88 (1-2)	accident	fundraiser
32	intelligence	health	activity	iq	1 (1-1)	being	being
33	sports	curricular	intramural	intramural	0 (1-2)	intramural	intramural
34	automobiles	solar	transaxle	carmaker	1 (1-2)	cars	cars
35	fitness	fashion	slang	beauty	1 (1-1)	beauty	beauty
36	climate	sustainability	mesoclimate	ecoclimatic	8 (1-2)	heat	ecoclimatic
37	politics	investment	portfolio	capital	2 (1-1)	portfolio	capital
38	mood	fitness	personality	regimen	1 (1-1)	fettle	fettle
39	cuisine	language	chana	spanish	5 (1-1)	italian	spanish
40	art	food	imagination	gastronomy	5 (1-1)	school	gastronomy
41	marriage	beauty	intimacy	love	1 (1-1)	queen	queen
42	violence	independence	peace	peace	1 (1-1)	war	war
43	fantasy	violence	chaos	chaos	31 (1-2)	chaos	chaos
44	criticism	law	eulogy	textualism	2 (1-1)	writing	writing
45	mining	himalayas	gulf	snoff	24 (1-2)	apex	apex
46	aviation	network	collection	backhaul	3 (1-1)	bug	codeshare
47	music	chore	polyvalence	burden	1 (1-1)	pleasure	pleasure
48	semiconductor	superconductor	semiconductor	magnetoresistance	1 (1-1)	conductor	conductor
49	coke	sucrose	carbonite	sugar	1 (1-1)	sugar	sugar
50	game	book	game	octavofinal	21 (1-1)	type	card
51	problems	election	test	issue	8 (1-2)	test	issue
52			15	21		24	30
53	No satisfactory word - 13 rows						

Figure 6.1: Analysis of methods for finding next suitable topic.

Words highlighted in purple are satisfactory words and in pink are medi-
umly satisfactory (because the intermediate words are related to start topic
and target topic directly with different context meaning).

Performance of fine-tuned DialoGPT

We assessed the performance of the fine-tuned DialoGPT model on a test dataset of size 552. The primary task of generating a response that contains the keyword given in input is considered successful when a word having a similarity score greater or equal to 0.9 is found in a conversation. The similarity score is obtained from WordNet [3].

We also calculated the BLEU score [4] and ROUGE Score [2] In addition, we assessed semantic similarity between reference utterances and generated responses using BERT embeddings, revealing a significant cosine similarity of 86% and indicating substantial alignment between the two.

Table 6.1: Fine-tuned DialoGPT across various metrics

Number of epochs	Success rate	ROUGE-L	BLEU ¹⁻⁴	Cosine similarity
1	70.29%	7.39%	27.79%	86.12%
2	60.14%	7.65%	27.70%	86.02%
3	54.53%	10.58%	29.08%	86.52%

From Table 6.1, it is evident that although the BLEU and ROUGE scores improved with the number of training epochs, the success rate of generating responses with the given input keyword declined. Since success rate is crucial for us, we opt for the first epoch, which has a success rate of 70.29%. (The BLEU scores presented in the table represent the full BLEU score that accounts for the overlap across 1-4 ngrams)

Some manually conducted chats with the chatbot

```
Target word: dance

User> What are you doing tonight?

current keyword: tonight
next keyword: music

TBot> i am listening to some metal music and relaxing. how about you?

User> I am going to watch a movie

current keyword: movie
next keyword: music

TBot> what are you watching?

User> I am going to watch Inception

current keyword: inception
next keyword: film

TBot> that is a good one. i like films that are suspenseful.

User> What type of films do you like to watch?

current keyword: film
next keyword: moonwalker

TBot> i like to see movies about moonwalker.

User> What is moonwalker

current keyword: moonwalker
next keyword: dance

TBot> it is a dance movie. i like to dance.
```

Figure 6.2: Successful chat 1.

```
Target word: vacation
User> hii buddy
current keyword: hii
next keyword: hii
TBot> hey, how are you?
User> i am great. are you free this evening lets party
current keyword: party
next keyword: home
TBot> i can. i am at home.
User> okay cool.
current keyword: home
next keyword: vacation
TBot> do you have any vacations coming up?
```

Figure 6.3: Successful chat 2.

```
Target word: sports
User> do you want some food
Current Keyword: food
Next Keyword: bowl
TBot> i do not really like food. i like to bowl.
User>
```

Figure 6.4: Chat with an abrupt transition.

In Figures 6.3 and 6.2, the target topic is achieved smoothly in 5 and 3 responses respectively. In Figure 6.4, we see an abrupt topic transition from food to bowl (sports). This is due to the fact that “bowl” is related to both food and sports but with different meanings. It can be a bowl of rice or bowl in cricket.

Observations in automated chats

We conducted 50 conversations between our topic steering chatbot and BlenderBot-400M-distill by randomly selecting the target topic and initiating the conversation. The conversation length was capped to 8 responses by our chatbot. Our chatbot successfully directed the conversation towards the target topic in 21 instances in an average of 3.52 responses of our chatbot.

Nevertheless, in many unsuccessful conversations, our chatbot identified the target topic as the next keyword for its response, but it failed to incorporate it effectively. An example chat is shown in Figure 6.5. This highlights the challenge we face: the low success rate of our fine-tuned DialoGPT in generating responses with the given keyword.

```
target> disco
user>> what do you do for living
current kw> do
next kw> music; method: 1st hop - 1st hop
Tbot>> i am a music teacher. i teach high school music.
Blender>> That's awesome! I bet you have a lot of fun teaching music. Do you enjoy it?
current kw> music
next kw> disco; method: 1st hop - 1st hop
Tbot>> i do, but i like the circus more. i also love to dance.
Blender>> I bet that's an interesting job. What kind of dancing do you like to do?
current kw> dancing
next kw> disco; method: 1st hop - 1st hop
Tbot>> i like to dance to country music, and i like to play in the circus.
```

Figure 6.5: Fine-tuned DialoGPT fails to generate response containing the next keyword.

Chapter 7

Conclusion

In this work, we tried to develop an effective target-oriented topic steering open-ended conversation system. Our proposed intersection method to find the next topic of response using ConceptNet reduced the search space and facilitated a seamless transition in conversation. Our chatbot was able to successfully steer the conversation towards the target topic when the current conversation topic and target topic were not very far in most cases. However, it also suffers in many cases due to the fact that fine-tuned DialoGPT fails to incorporate the given input keyword in generated response effectively. Employing a larger model can enhance the fluency of responses and increase the likelihood of successfully generating a reply based on the given input keyword.

Bibliography

- [1] Yosuke Kishinami, Reina Akama, Shiki Sato, Ryoko Tokuhisa, Jun Suzuki, and Kentaro Inui. Target-guided open-domain conversation planning. *arXiv preprint arXiv:2209.09746*, 2022.
- [2] Chin-Yew Lin. Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*, pages 74–81, 2004.
- [3] George A Miller. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41, 1995.
- [4] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. Bleu: a method for automatic evaluation of machine translation. In Pierre Isabelle, Eugene Charniak, and Dekang Lin, editors, *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 311–318, Philadelphia, Pennsylvania, USA, July 2002. Association for Computational Linguistics.
- [5] Jinghui Qin, Zheng Ye, Jianheng Tang, and Xiaodan Liang. Dynamic knowledge routing network for target-guided open-domain conversation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 8657–8664, 2020.
- [6] Robyn Speer, Joshua Chin, and Catherine Havasi. Conceptnet 5.5: An open multilingual graph of general knowledge. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, 2017.

- [7] Jianheng Tang, Tiancheng Zhao, Chenyan Xiong, Xiaodan Liang, Eric P. Xing, and Zhiting Hu. Target-guided open-domain conversation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020.
- [8] Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. Personalizing dialogue agents: I have a dog, do you have pets too? *arXiv preprint arXiv:1801.07243*, 2018.