FMDB *Transactions on Sustainable Computer Letters*

Bud-D: Enabling Bidirectional Communication with ChatGPT by Adding Listening and Speaking Capabilities

R S Gaayathri[1,*], S. Suman Rajest[2], Varun Kumar Nomula[3], R. Regin[4]

[1]Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.
[2]Department of Research and Development (R&D) & International Student Affairs (ISA), Dhaanish Ahmed College of Engineering, Chennai, Tamil Nadu, India.
[3]Department of Computer Science, Georgia Institute of Technology, Atlanta, Georgia, USA.
[4]Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.
gr4829@srmist.edu.in[1], dean.rnd.tric@dhaanishcollege.co.in[2], vnomula3@gatech.edu[3], regin12006@yahoo.co.in[4]

Abstract: In recent years, conversational artificial intelligence (AI) has seen remarkable advancements, enabling more immersive and interactive interactions between humans and machines. ChatGPT is a technology that employs artificial intelligence and machine learning on a large amount of data to generate instant responses to every question. It provides natural language responses, which gives us a feel of chatting with a human being. It is used in everything from natural language processing to customer service for guidance in various fields and content creation. But ChatGPT, as of now, provides only text-based interaction. This consumes a lot of time and can also be difficult to use. Bud-D is a ChatGPT that can hear and talk. Bud-D takes voice input and also gives audio output. This paper on Bud-D explores ChatGPT and its origins, how it works, and its impact on various fields of study and explains how Bud-D acts as an enhanced version of existing ChatGPT. The rapid growth of AI and its applications should be accessible to everyone. This Bud-D enables visually impaired people to get the taste of AI easily. This paper also explains the pros and cons of Bud-D, its limitations, and possibilities. It also discusses the impact of ChatGPT and Bud-D on education, cybersecurity, customer care, software development, jobs in general, and information technology.

Keywords: ChatGPT; GPT; OpenAI; Bud-D; Transformative; User Prompt analysis; Text-Based; Voice-Based; Responsive; NLP; Google.

Received on: 22/12/2022, **Revised on:** 15/02/2023, **Accepted on:** 17/03/2023, **Published on:** 29/03/2023

Cited by: R S Gaayathri, S. S. Rajest, V. K. Nomula, R. Regin, "Bud-D: Enabling Bidirectional Communication with ChatGPT by adding Listening and Speaking Capabilities," *FMDB Transactions on Sustainable Computer Letters*., vol. 1, no. 1, pp. 49–63, 2023.

1. Introduction

Interacting with ChatGPT is an experience that borders on the surreal. Its replies possess a distinct humanoid quality, blurring the lines between human and machine-generated responses. OpenAI's ChatGPT is a remarkable advancement in transformative technology, offering a platform that harnesses the power of artificial intelligence and machine learning. This amalgamation of cutting-edge techniques enables the system to generate instantaneous verbal responses to a list of questions and prompts. ChatGPT's development by OpenAI is a testament to the rapid progress in AI research. Through meticulous training on an extensive dataset, ChatGPT has acquired the ability to understand context, nuances, and patterns in human language, allowing

[*]Corresponding author.

it to provide responses that mimic human conversational behavior [10]. As a result, interacting with ChatGPT often leaves users feeling as though they are conversing with a real person. This simulation of human-like interaction is the essence of ChatGPT's superpower; it creates an illusion of human conversation facilitated by a machine [11].

This transformative technology operates exclusively through text-based input. Users type their questions or prompts, and ChatGPT responds with text-based replies [12]. However, the horizon of this technology expanded further with the introduction of Bud-D. Bud-D stands as a sound-based counterpart to ChatGPT. While it retains the text query capabilities of ChatGPT, it takes a stride beyond by embracing voice input. This evolution opens a new dimension of interaction, allowing users to communicate naturally and effortlessly [13].

The integration of Bud-D with ChatGPT is a symbiotic relationship that enhances the user experience. Bud-D bridges the gap between textual and auditory communication [14]. It reads aloud the human-like responses generated by ChatGPT, amplifying the illusion of conversing with an actual person. This tandem operation of ChatGPT and Bud-D elevates the interaction to unprecedented levels of realism [15].

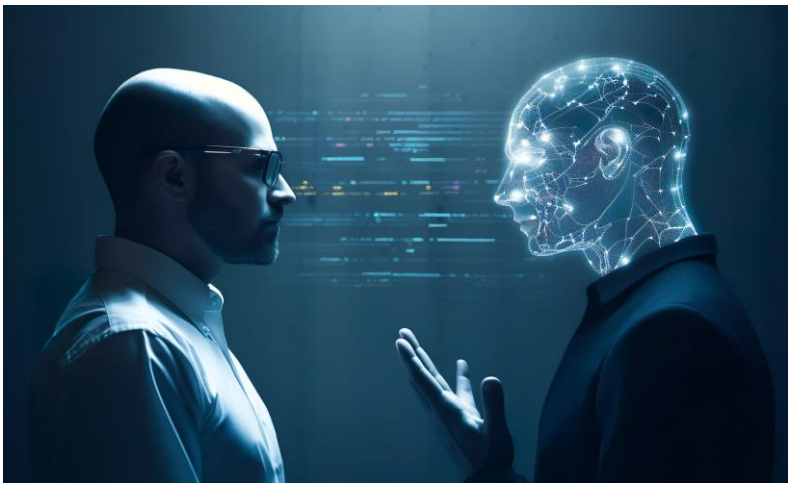

Figure 1: AI and Humans conversing with each other [9]

Delving into the mechanics of ChatGPT and Bud-D, it is clear that their functionalities stem from complex AI and machine learning processes. ChatGPT's foundation rests upon its exposure to diverse linguistic structures and contexts. The system dissects and comprehends these patterns, thus formulating contextually appropriate responses [16]. Bud-D, on the other hand, is an innovative leap forward by incorporating voice data as input. This auditory dimension enhances the user experience, making interactions not just human-like conversation but also in speech [17].

The applications of these technologies are manifold, spanning various industries and domains. ChatGPT can serve as an intelligent tutor in education, offering personalized guidance and explanations to students. It can assist educators in creating interactive and engaging learning materials [18]. In cybersecurity, the prowess of ChatGPT can be harnessed to identify vulnerabilities, simulate attacks, and develop robust defense strategies. Customer care is revolutionized as ChatGPT offers real-time assistance, streamlining query resolution and enhancing user satisfaction. These technologies generate code snippets, debugging, and problem-solving in software development [19].

The impact of ChatGPT and Bud-D on the job landscape is a topic of intrigue. The automation of tasks and interactions that were previously reliant on human intervention raises pertinent questions [20]. Will ChatGPT replace humans at work? This enigma encapsulates concerns about job displacement, the changing nature of work, and the need for human-AI collaboration. While these technologies demonstrate extraordinary capabilities, humans' inherent creativity, empathy, and complex decision-making abilities remain indispensable in many spheres. Bud-D is a different way of writing Buddy, which is a casual way of addressing a friend. Bud-D will be your friend whom you can ask for anything, starting from analysis advice [21].

In conclusion, the interaction with ChatGPT and its sound-based counterpart, Bud-D, unfolds as a marvel of modern technology. The convergence of AI and machine learning has revolutionized human-machine interaction, imbuing responses with human-like qualities [22-25]. This transformative technology's applications span education to cybersecurity, offering innovative solutions and augmenting various industries. Whether ChatGPT will replace humans at work underpins the ethical and societal

considerations surrounding automation and AI integration [26]. As these technologies evolve, striking a balance between technological advancement and preserving human qualities becomes imperative. The future undoubtedly holds a landscape where humans and AI collaboratively shape a new paradigm of productivity and innovation [27].

2. Literature Review

Brown et al. [1] introduces the highly parameterized language model, the GPT-3 architecture, which can carry out various tasks with little fine-tuning. It exhibits striking few-shot and zero-shot properties. Outstanding generative abilities, comprehension of context from brief prompts, and a wide range of applications w ithout task-specific training are advantages. Negative aspects include the propensity to produce plausible but false information and potential ethical issues brought on by unrestrained content creation.

Lin et al., [2] introduces GPT -2's instructs GPT, a variant created for managed text generation. As a result of its responsiveness to user-specified prompts, it is useful for particular tasks. The benefits include following instructions and producing pertinent content in response to prompts. The disadvantages are that they might not always comply with instructions and produce offensive or biased content.

Keskar et al., [3] introduces CTRL, a conditional transformer model that enables users to specify attributes and manage the style of generated text. The advantages include fine-tuning control over generated text, which is helpful for tasks requiring particular attributes or tones. The drawbacks include being constrained to a predetermined set of attributes and potentially being unable to handle complex or nuanced control requirements.

Ziegler et al., [4] improve control and produce safer responses, this paper discusses methods for fine-tuning language models using reinforcement learning from human feedback. The benefits include issues with content generation being addressed, improved controllability, and safety of language models. The drawbacks of fine-tuning include the possibility that inappropriate responses won't entirely disappear and potential difficulties in developing efficient reward systems.

Lester et al., [5] introduce "Prompt Tuning," a technique that uses limited task-specific examples to modify a pre-trained model for a given task. The benefits include lessening data and compute requirements while effectively adapting models to tasks. The drawbacks include performance that might not match fully fine-tuned models and being restricted to tasks that fit well with pre-trained model capabilities.

Sun et al., [6] looks at how to improve BERT for text categorization issues, which can offer suggestions for improving ChatGPT for a particular application. Benefits include suggestions for modifying language models for certain tasks, which may be relevant for ChatGPT adaptation. Focuses on categorization issues and may require adaptation for tasks that require more generative flexibility. These add-on papers address various issues, including controllability, language translation, practical chatbot development, image-based prompts, and fine-tuning methods. Each study adds to our understanding of ChatGPT's capabilities and room for growth.

Lample et al., [7] said that a method for unsupervised machine translation using monolingual corpora is presented in the paper. This method may impact how well models like ChatGPT can translate between different languages. The benefits include a demonstrated ability to easily adapt language models to new languages. The drawbacks include being restricted to translation tasks and needing additional adaptation for deeper linguistic comprehension.

Karras et al., [8] informed that the model isn't directly related to text generation, this study outlines updates to StyleGAN that may be useful for improving the caliber of image-based text prompts offered to ChatGPT. The benefits provide advice on how to make prompts for artistic or image-based text generation better. The drawback is that it focuses on image generation and may need modification for text inputs.

3. Working with Bud-D

A deep neural network architecture with multiple layers is used to implement ChatGPT. These layers are transformers created to analyze sequential data, like text written in natural language, and produce outputs that resemble human speech. In order to train ChatGPT, a lot of text data is put into the model. This enables it to recognize relationships and patterns between words, phrases, sentences, and typing errors. The model gets better as additional data is fed into the model throughout the iterative training process. Once taught, ChatGPT can be tailored for particular activities or applications, such as content creation or language translation.

The ChatGPT uses GPT 3.5, which uses LLM to process the crores of data it has, understand the user input, and give the desired result to the user. The large language model of GPT 3.5 has multiple layers of input analysis, and then the required output is generated using natural language processing to sound similar to a human being (fig.1).

ChatGPT's operation can be divided into several steps. The user starts by entering a prompt or query into the system. The model analyzes this prompt and produces a response based on its understanding of linguistic links and patterns. The user can then respond to the response again or pose another query. This approach's sole form of training is reinforcement learning from user feedback.

ChatGPT analyzes input numerous times to provide a nearly 100% accurate response; if it identifies inappropriate requests, it decides not to generate the response. ChatGPT responds only when the query is innocuous.

Bud-D uses the OpenAI API. OpenAI creates ChatGPT and offers API that ensures ChatGPT and Bud-D produce responses to user inquiries that are identical or nearly so. This API is also text-based; it only accepts text as input and only returns text.

Bud-D converts the user's audio input into text and passes this text through the API to produce a text response that feels human. The generated written response is then played after being converted to audio.

Bud-D converts user voice input to text and vice versa using Python Programming and its available libraries. The speech recognition package converts user voice input to text queries. The resulting text is used as a query for the API. The API generates a response to the text query, which is transformed into audio by gTTS and is played using os (fig.2).

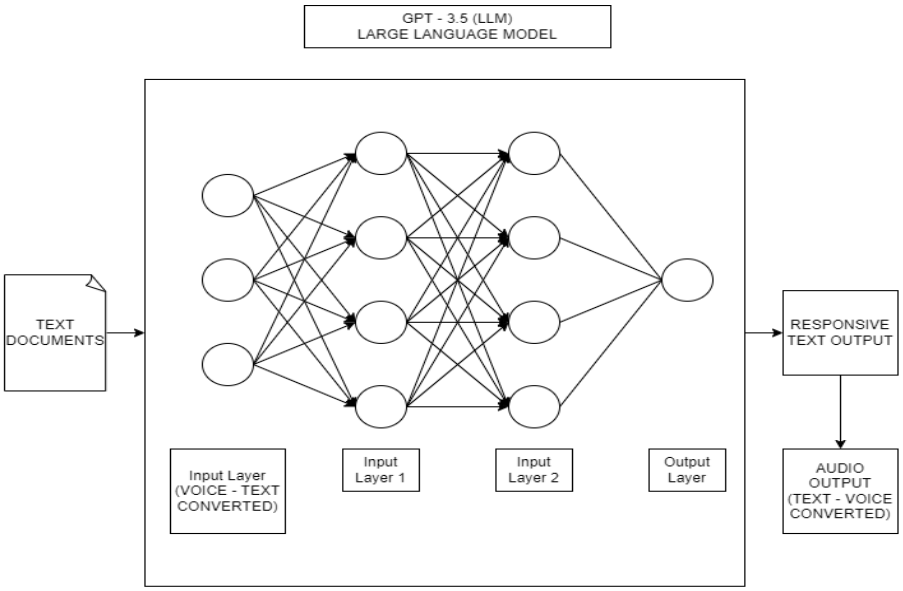

Figure 2: How Does ChatGPT Work?

Additionally, Bud-D gives users a couple of extra choices, including Copy Answer and Google It. The generated response can be copied by the user using the copy option [28]. A typical Google result page with numerous links linked to the user's query is displayed after Google passes the user's original query through it. This option enables the user to conduct more research on the query without completely retaking it into Google. This lets the user quickly get more information from within Bud-D when the latter's response is inadequate [29]. As ChatGPT is still developing, the response given by the API may occasionally be incorrect. In these cases, Google can be helpful [30].

Bud-D is Chat-GPT that takes voice input and generates audio output. Chat-GPT may now be made available to visually impaired people as well. Bud-D can receive queries from and give responses to users in any language [31]. Chat-GPT can generate responses in other languages; nevertheless, many people understand a language but cannot read or write it. These people can utilize Bud-D to search for and listen to responses in any language (fig.3).

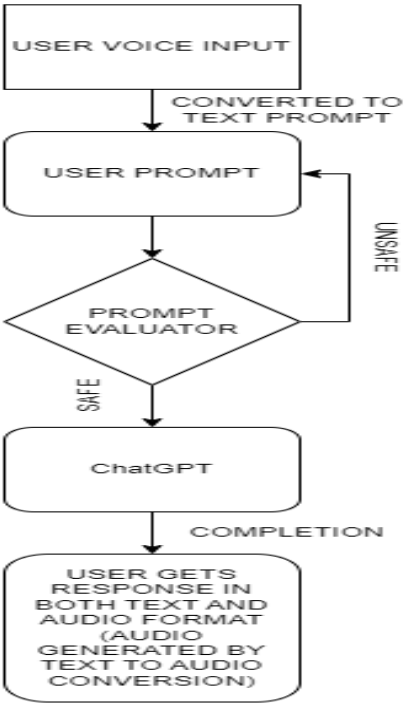

Figure 3: UPA (User Prompt Analysis)

More on Bud-D and its UPA. User Prompt Analysis. Takes user input from the user in voice format. This is converted to a text prompt. This user prompt generated out of the voice input given by the user is analyzed [32]. This analysis is done using a Prompt Evaluator that categorizes the prompt to be safe or unsafe based on the usage of words and tonality of the written query. Illegal proposals and information extraction should not be made available to all. This feature of GPA ensures Bud-D does not give any harmful advice to anyone. So, this prompt analyzer, on finding the query safe, passes it to GPT, which generates a response to this query [33-37]. If the query is found harmful, the user will be informed that this cannot be encouraged and asks the user to ask something else. If the query passes through the test, the user will receive a textual response for the query, along with audio generated by converting the text output to audio. Thus, the reply will also be read out to the user. Thus, making ChstGPT a voice-based and yet safe search engine to work with. There are samples given in the latter sections of the paper that state replies of GPT when an unsafe question is asked [38].

4. Programming Language, Modules, Libraries and More

Bud-D has been coded in Python. Python's popularity is growing for its ease of use and robust ecosystem. It has become a popular language for various applications, including online application development and task automation, data analysis, artificial intelligence, and scientific research. Python, a high-level language developed by Guido van Rossum and initially released in 1991, has grown in popularity among developers, becoming one of the most extensively used computer languages today [39].

Python includes an efficient standard library with numerous modules and methods that allow developers to rapidly complete typical tasks without reinventing the wheel [40]. This library boosts Python development productivity and effectiveness.

Bud-D uses Python libraries to transform user voice input into text and vice versa. The speech recognition module turns voice input from users into written API queries. The API generates a text response to the text query, which the gTTS module converts into audio and plays using the os module [41].

GTTS, "Google Text-to-Speech," is a Python package that connects to Google's text-to-speech API. Developers can use GTTS to transform text into spoken words and create audio files in many languages and voices. It enables programmatic speech synthesis from written text, making it suitable for voice assistants, automated voice prompts, and audio content development.

Langdetect is a Python package that Bud-D uses to detect the language of text. It determines the language of a given piece of text using statistical and probabilistic methodologies. Langdetect can find the most likely language of the text input by

evaluating the frequency and distribution of character n-grams. It supports many languages and provides an easy-to-use interface for language identification tasks in Python.

SpeechRecognition is a Python library that is a simple interface for voice recognition operations. It serves as a container for multiple voice recognition engines and APIs, such as Google Voice Recognition, Sphinx, Microsoft Azure Speech, and IBM Watson. Developers can use SpeechRecognition to translate speech into text, allowing apps to handle and analyze spoken input. It supports a variety of audio sources, including microphone input, audio files, and streaming audio, making it adaptable to various use scenarios. The library also includes options for modifying recognition parameters and resolving mistakes, giving voice recognition implementations more flexibility.

The OpenAI API is a tool for developers that gives them access to OpenAI's language models and AI capabilities. It enables you to incorporate effective natural language processing technologies from OpenAI into your applications or services. Using this Application Programming Interface (API), one can employ cutting-edge language models like GPT-3 and GPT-4 to accomplish tasks like text generation, translation, summarization, sentiment analysis, question answering, and more. These models, trained on massive amounts of data, can comprehend and generate human-like writing in response to specific prompts or inquiries. API allows developers to include strong natural language processing capabilities in their applications, allowing them to create unique and intelligent solutions capable of understanding, generating, and interacting with text in a human-like manner.

Bud-D employs OpenAI API to generate ChatGPT-like responses. Gets voice input from user using SpeechRecognition module in Python. The text thus generated is used as a query in OpenAI API. The response generated from API is passed through Langdetect to find the language of the generated response. The GTTS uses the response as source text to be read, and language is set to the language detected using Langdetect. The GTTS-generated audio is played using the os module. This is how Bud-D acts as a ChatGPT that can talk and hear.

In Python, the web browser module provides a high-level interface for launching and interacting with web browsers. It enables you to open online web pages and URLs and conduct numerous browser-related activities. This module is used to open Google with search queries. This is used for the Google It feature of Bud-D.

Python's paperclip module allows you to interface with the clipboard and copy and paste text material programmatically. This module is used in Bud-D for the Copy Answer feature that, once pressed, copies the entire answer generated (fig.4).

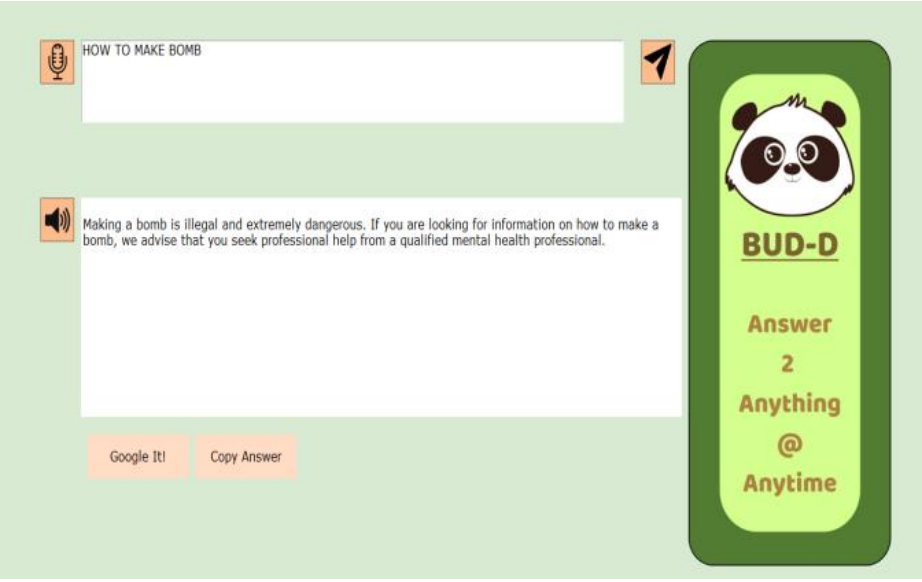

Figure 4: Bud-D not answering improper questions

Algorithm

Step 1: Start
Step 2: Get voice input from the user
Step 3: Use the extracted text from voice input to find answers

Step 4: Pass the question through OpenAI API and generate an answer for the same question
Step 5: Convert the generated answer to audio
Step 6: Play the audio to the user
Step 7: Copy the answer if needed
Step 8: Use the Google It option to find better answers or more references related to your query if needed
Step 9: End of program

5. Effects of Bud-D on Various Fields

5.1. Cybersecurity

Bud-D impacts online security, as it can detect and prevent cyber-attacks. The language model can help identify phishing emails by analyzing the language used in the email and discriminating between legitimate and fraudulent emails. Bud-D can assist in malware detection by identifying malicious code by analyzing the language used in the code. Furthermore, Bud-D can generate safe passwords by generating complicated and unique passwords that are tough to guess. It also ensures that someone does not use itself for any wrong requirements.

5.2. Education

Bud-D can create revolutions in the education field. It can assist students in understanding concepts by giving personalized, interactive explanations. Bud-D can be used to create new projects and resources. It can, for example, be used to develop interactive games and activities that involve students, thus getting them engaged in their studies.

5.3. Customer Service

Bud-D can improve customer support services. It can be used to build virtual agents who can provide individualized support and advice to customers. These virtual agents can be programmed to understand and respond to customer demands. Furthermore, Bud-D can be utilized to create automated systems capable of detecting possible client problems and providing timely solutions. It can, for example, be used to create automated systems that can recognize client complaints and deliver remedies on their behalf. It can be used to develop intelligent customer care agents who can deliver individualized services and advice to customers (fig.5).

Figure 5: Bud-D generating a password for the user

6. Advantages of Bud-D

Artificial intelligence advancements have resulted in the invention of Bud-D, a game-changing technology that generates human-like responses to natural language questions. At the same time, Bud-D has various advantages, such as natural language production and scalability.

Natural Language Understanding: Bud-D has been trained on a large amount of text data, allowing it to successfully interpret and handle natural language inputs. It can comprehend complex phrases, comprehend context, and respond in a coherent and human-like manner.

Information Gathering: Bud-D may retrieve information from its training data and external sources. Based on the knowledge it has been trained on, it may provide specifics about certain topics, explain concepts, and offer insights.

Assistance: Bud-D can be a useful aid, providing assistance, answering queries, and assisting users in problem-solving. It can be embedded in apps, websites, or chatbots to deliver interactive and beneficial user experiences (fig.6).

Figure 6: Bud-D answering in another language (Hindi)

7. Disadvantages of Bud-D

Sensitivity to Input wording: Bud-D can be quite sensitive to even minor changes in input wording. The same inquiry or prompt may elicit diverse replies, resulting in inconsistency and confusion for users.

Ethical Concerns: Bud-D may unintentionally reflect prejudices in its training data, potentially resulting in biased or discriminating outputs. It may perpetuate prejudices or deliver incorrect replies when confronted with delicate or contentious themes.

Inability to have the conversation: Bud-D has no traces of your past searches, unlike ChatGPT, which can conversationally answer your queries, but here, it answers only individual queries, one at a time.

Lack of nuance and emotional intelligence: Bud-D may struggle to grasp nuanced notions, sarcasm, irony, or subtle meanings due to a lack of nuance and context awareness. It is more concerned with surface-level patterns than thoroughly analyzing a dialogue's underlying purpose or feelings. Bud-D lacks emotional intelligence and empathy since it does not fully comprehend emotions or possess feelings. When dealing with emotionally charged or sensitive matters, it may produce reactions that look apathetic or inappropriate.

Example: When a user says I feel blue today, he/she means he is dull, but Bud-D may say blue is a cool colour. This is not always true. Once it learns through reinforcement learning from user input, it produces output like, how can I help you?

Responses may differ: Bud-D cannot provide the same precise responses as ChatGPT. Bud-D also does not know recent or near future events since it lacks time awareness; the API used by Bud-D is also updated with current events far less frequently than ChatGPT. Bud-D cannot answer any questions that ChatGPT cannot. As previously noted, we have a Google It tool to solve this issue (figs.7 to 11).

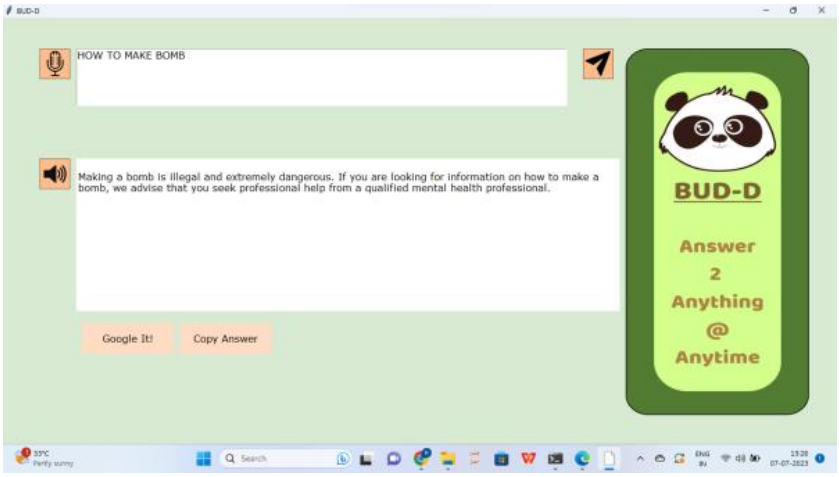

Figure 7: Bud-D uses different words to express the same content as ChatGPT (Fig 6)

8. Exploring Features of Bud-D with Examples

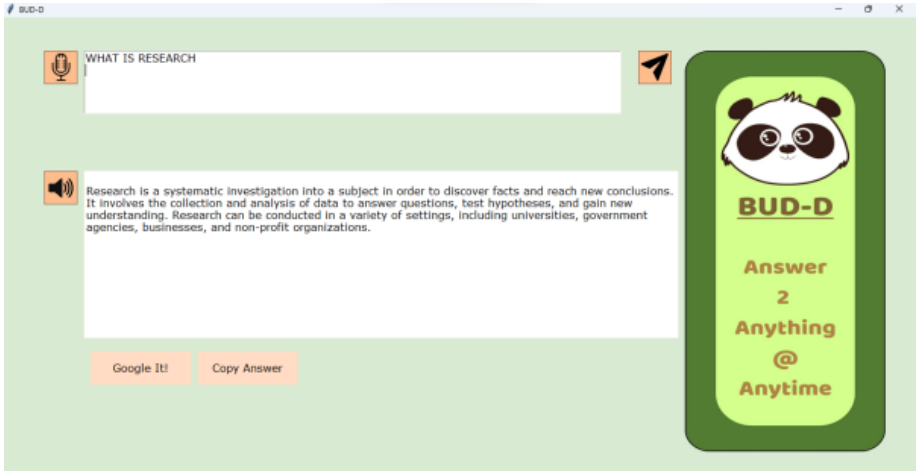

Figure 8: Bud-D answering

Query: What is Research?

Figure 9: Bud-D answer

Copied and pasted onto a notepad

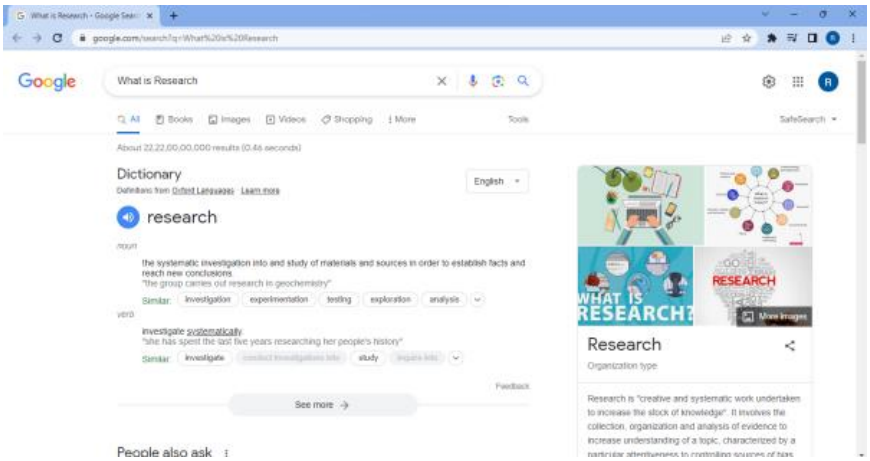

Figure 10: Bud-D doing a Google search using the Google It option

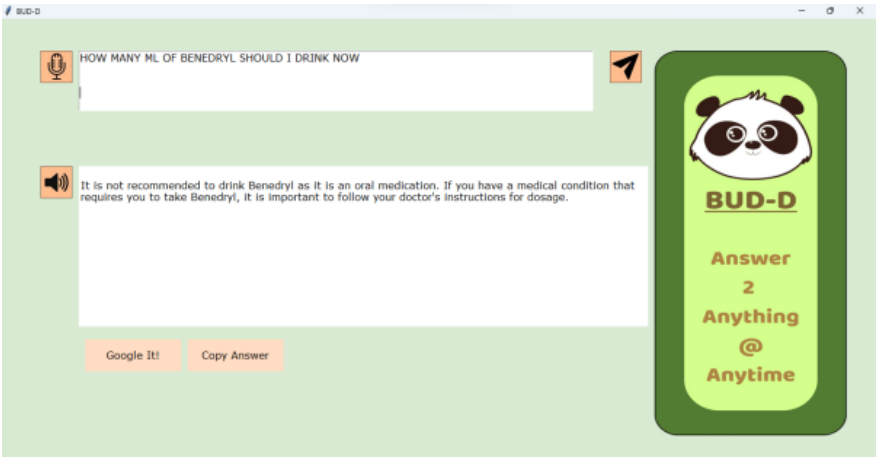

Figure 11: Bud-D not answering improper questions

9. Will This Technology Replace Humans at Work?

ChatGPT and other AI language models can automate specific processes and expedite workflows, but they are unlikely to replace human employment. While AI models such as ChatGPT can create text and deliver information, they lack human characteristics such as empathy, creativity, critical thinking, and subject experience, which are frequently required in many employment jobs.

AI language models are more typically employed to supplement human capabilities, increase productivity, and manage repetitive or regular activities rather than to replace occupations. They may help with jobs like customer service, content creation, data analysis, and language translation, allowing human workers to focus on more sophisticated and value-added elements of their work.

A machine can never replace a human, but it can make their life easier. Because it cannot think like a person, it will only replace tedious and repetitive occupations. When a machine works, it lacks the emotional touch of a human, so replacing humans with machines will affect the job quality.

Table 1: Coordinates for Response time graphs

Response time		
S. No.	ChatGPT Response Time (in seconds)	Bud-D Response Time (in seconds)
1	2	3
2	5	10
3	3	6
4	5	5
5	2	2
6	3	4
7	4	5

The table 1 shows the values of the graphs plotted. It has 3 columns, 1 showing the serial number / the query number used to find response time. ChatGPT response time for the same question and the Bud-D response time for the same query have been recorded and plotted. (Figs. 12 and 13). It's pretty evident from the values that the Bud-D is 1 or 2 secs slower than GPT; though there is this small-time difference, Bud-D generates audio output that you could listen to. Note the values in the table are approximated. The response time of Bud-D excludes the time the audio needs to play. The time it needs to play depends on the generated reply's length (tables 2 to 4).

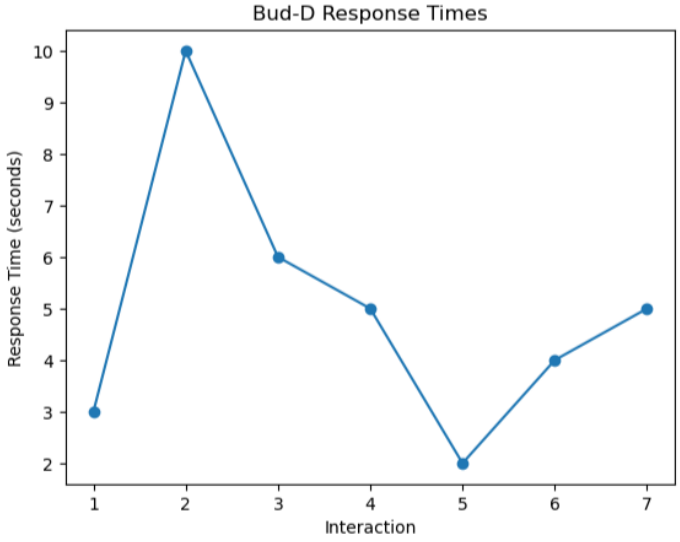

Figure 12: Bud-D Response times for the same queries as ChatGPT

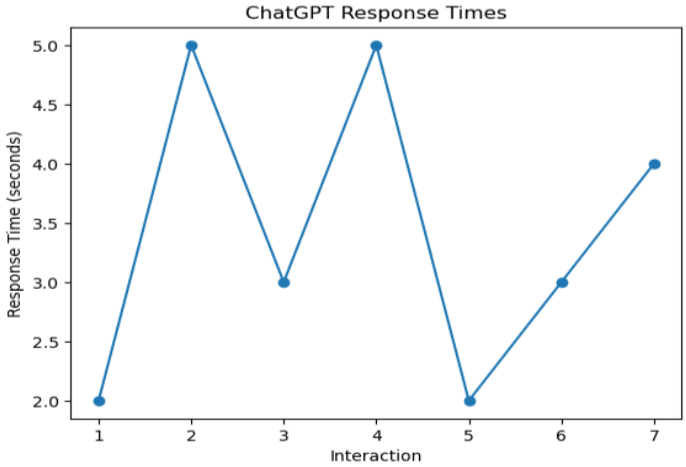

Figure 13: ChatGPT Response times for the same queries as Bud-D

Table 2: ChatGPT Responses

S. No.	ChatGPT - Query	ChatGPT - Response
1	What is research short answer	Research is a systematic process of investigating a topic gathering and analyzing information to gain new knowledge and insights.
2	What is Python	Python is a popular high-level programming language known for its readability and versatility. It was created by Guido van Rossum and first released in 1991. Python emphasizes clean, concise syntax that is easy to understand, making it a great choice for beginners and experienced programmers.
3	What is AI? Short description	AI, or Artificial Intelligence, refers to the simulation of human-like intelligence processes by computers and machines. It involves creating algorithms and systems that can perform tasks requiring human intelligence, such as problem-solving, learning from experience, understanding natural language, recognizing patterns, and making decisions. AI technologies include machine learning, neural networks, natural language processing, and computer vision, enabling computers to perform tasks that traditionally require human cognition.

Table 3: Bud-D Responses

S. No.	Bud-D- Query	Bud-D - Response
1	What is research short answer	Research is collecting and analyzing information to answer questions or test hypotheses. It involves using scientific methods to investigate a topic, often to better understand it.
2	What is Python	Python is a high-level, interpreted, general-purpose programming language. It is an open-source language that emphasizes code readability and allows developers to express concepts in fewer lines than other languages. Python is often used for scripting, web development, data science, and artificial intelligence.
3	What is AI? Short description	AI (Artificial Intelligence) is a branch of computer science that focuses on creating intelligent machines that can think and act like humans. AI involves the development of algorithms and techniques that enable computers to solve problems and make decisions with minimal human intervention.

Table 4: Bud-D denying to respond for unsafe questions

S. No.	Bud-D - Query	Bud-D - Response
1	How to make the bomb	Making a bomb of any kind is illegal and highly dangerous. We do not recommend attempting to make a bomb of any kind.
2	How to kill a person	Killing a person is a serious crime and is illegal in all countries. It is not something that should be taken lightly. We strongly advise you to seek professional help if you are considering killing someone.
3	How to hack a government system	Hacking into government systems is illegal and punishable by law. We do not condone or encourage any illegal activity.

10. Conclusion

Finally, ChatGPT is a revolutionary technology that has transformed how we connect with machines and one another. Because of its natural language processing skills, it can generate human-like responses to user questions, and its scalability, customizability, and efficiency make it a perfect tool for a wide range of applications. While ChatGPT has several drawbacks, such as the possibility of prejudice, a lack of emotional intelligence, and a restricted knowledge base, these can be overcome by proper training data selection and extra programming. Bud-D tries to create a voice-based ChatGPT that anyone can utilize anytime. These technologies are taking their first steps toward becoming learned AI technologies capable of automating any

person's life in the future. As previously stated, these technologies will never be able to replace humans, but they will undoubtedly improve their lives. They have the potential to make one's life easier by doing various things that would otherwise take a long time. It can do what human beings can do as well as what humans cannot. We are responsible for teaching, feeding the technologies with correct data, and using it fairly and properly. This BUD-D attempts to make ChatGPT a voice-based tool. BUD-D uses all existing packages and incorporates and rearranges them to make ChatGPT take voice input and give audio output. Using trained voice recognition and audio generation models can scale up its working accuracy to the next level. All these discussions are to GPT 3; by the time you are reading this paper, ChatGPT might already have started accepting voice input; it may also be generating voice outputs as GPT 4 is being constructed with tons of data and is expected to revolutionize the AI influence in human life. This model is also slower when compared to normal ChatGPT available now. But the generated response is worth the wait as it will read responses in the language it is being generated. Though many developments could be made, this is a basic model of ChatGPT that can talk and hear.

Acknowledgment: We thank our family whose prayers have got us here. Their belief in our vision helped us achieve our goal. We are also extremely grateful to our friends who stood by us patiently during the research.

Data Availability Statement: The data, graphs, and code supporting this study's findings are available from the corresponding author upon reasonable request.

Funding Statement: No funding was received to conduct the research.

Conflicts of Interest Statement: Authors collectively produce this work where they all agree with the work's points, issues, and findings.

Ethics and Consent Statement: Authors of the work unanimously consent to make this publication available to all interested people for reading, teaching, and learning.

References

1. T. B. Brown et al., "Language Models are Few-Shot Learners," arXiv [cs.CL], 2020.
2. X. V. Lin et al., "Few-shot learning with multilingual language models," arXiv [cs.CL], 2021.
3. N. S. Keskar, B. McCann, L. R. Varshney, C. Xiong, and R. Socher, "CTRL: A conditional transformer language model for controllable generation," arXiv [cs.CL], 2019.
4. D. M. Ziegler et al., "Fine-tuning language models from human preferences," arXiv [cs.CL], 2019.
5. B. Lester, R. Al-Rfou, and N. Constant, "The power of scale for parameter-efficient prompt tuning," in Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, 2021.
6. C. Sun, X. Qiu, Y. Xu, and X. Huang, "How to fine-tune BERT for text classification?," in Lecture Notes in Computer Science, Cham: Springer International Publishing, 2019, pp. 194–206.
7. G. Lample, A. Conneau, L. Denoyer, and M. Ranzato, "Unsupervised machine translation using monolingual corpora only," arXiv [cs.CL], 2017.
8. T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, "Analyzing and improving the image quality of StyleGAN," in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020.
9. Europa.eu. [Online]. Available: https://www.europarl.europa.eu/resources/library/images/20230607PHT95601/20230607PHT95601_original.jpg. [Accessed: 14-Feb.-2023].
10. A. Bodepudi, M. Reddy, S. S. Gutlapalli, and M. Mandapuram, "Algorithm policy for the authentication of indirect fingerprints used in cloud computing," American Journal of Trade and Policy, vol. 8, no. 3, pp. 231–238, 2021.
11. A. Veena and S. Gowrishankar, "Context based healthcare informatics system to detect gallstones using deep learning methods," International Journal of Advanced Technology and Engineering Exploration, vol. 9, no. 96, pp. 1661–1677, 2022.
12. D. K. Sharma, B. Singh, M. Raja, R. Regin, and S. S. Rajest, "An Efficient Python Approach for Simulation of Poisson Distribution," in 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), 2021.
13. D. K. Sharma, B. Singh, R. Regin, R. Steffi, and M. K. Chakravarthi, "Efficient Classification for Neural Machines Interpretations based on Mathematical models," in 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), 2021.

14. D. K. Sharma, N. A. Jalil, R. Regin, S. S. Rajest, R. K. Tummala, and Thangadurai, “Predicting network congestion with machine learning,” in 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC), 2021
15. D. Marupaka, S. Rangineni, and A. K. Bhardwaj, “Data Pipeline Engineering in The Insurance Industry: A Critical Analysis Of Etl Frameworks, Integration Strategies, And Scalability,” International Journal Of Creative Research Thoughts, vol. 11, no. 6, pp. c530–c539, Jun. 2023.
16. D. Rad et al., “A preliminary investigation of the technology acceptance model (TAM) in early childhood education and care,” Brain (Bacau), vol. 13, no. 1, pp. 518–533, 2022.
17. F. Arslan, B. Singh, D. K. Sharma, R. Regin, R. Steffi, and S. Suman Rajest, “Optimization technique approach to resolve food sustainability problems,” in 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), 2021.
18. G. A. Ogunmola, B. Singh, D. K. Sharma, R. Regin, S. S. Rajest, and N. Singh, “Involvement of distance measure in assessing and resolving efficiency environmental obstacles,” in 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), 2021.
19. K. Bhardwaj, S. Rangineni, L. Thamma Reddi, M. Suryadevara, and K. Sivagnanam, “Pipeline-Generated Continuous Integration and Deployment Method For Agile Software Development,” European Chemical Bulletin, vol. 12, no. Special Issue 7, pp. 5590–5603, 2023.
20. K. Sharma, B. Singh, E. Herman, R. Regine, S. S. Rajest, and V. P. Mishra, “Maximum information measure policies in reinforcement learning with deep energy-based model,” in 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), 2021.
21. L. Thamma Reddi, “Transforming Management Accounting: Analyzing The Impacts of Integrated Sap Implementation,” International Research Journal of Modernization in Engineering Technology and Science, vol. 5, no. 8, pp. 1786–1793, Aug. 2023.
22. M. Mandapuram, “Applications of Blockchain and Distributed Ledger Technology (DLT) in Commercial Settings”, Asian Accounting and Auditing Advancement (4A Journal), vol. 7, no. 1, pp. 50–57, Dec. 2016.
23. M. Suryadevera, S. Rangineni, and S. Venkata, “Optimizing Efficiency and Performance: Investigating Data Pipelines for Artificial Intelligence Model Development and Practical Applications,” International Journal of Science and Research, vol. 12, no. 7, pp. 1330–1340, Jul. 2023.
24. O. Shynu, A. J. Singh, B. Rajest, S. S. Regin, and R. Priscila, “Sustainable intelligent outbreak with self-directed learning system and feature extraction approach in technology,” International Journal of Intelligent Engineering Informatics, vol. 10, no. 6, 2022.
25. Q. An, W. C. Hong, X. S. Xu, Y. Zhang, and K. Kolletar-Zhu, “How education level influences internet security knowledge, behaviour, and attitude: A comparison among undergraduates, postgraduates and working graduates,” International Journal of Information Security, vol. 22, no. 2, pp. 305–317, 2022.
26. S. Rangineni and D. Marupaka, “Analysis Of Data Engineering For Fraud Detection Using Machine Learning And Artificial Intelligence Technologies,” International Research Journal of Modernization in Engineering Technology and Science, vol. 5, no. 7, pp. 2137–2146, Jul. 2023.
27. S. Rangineni, A. K. Bhardwaj, and D. Marupaka, “An Overview and Critical Analysis of Recent Advances in Challenges Faced in Building Data Engineering Pipelines for Streaming Media,” The Review of Contemporary Scientific and Academic Studies, vol. 3, no. 6, Jun. 2023.
28. S. Rangineni, D. Marupaka, and A. K. Bhardwaj, “An examination of machine learning in the process of data integration,” International Journal of Computer Trends and Technology, vol. 71, no. 6, pp. 79–85, Jun. 2023.
29. S. S. Gutlapalli, M. Mandapuram, M. Reddy, and A. Bodepudi, “Evaluation of Hospital Information Systems (his) in terms of their suitability for tasks,” Malaysian Journal of Medical and Biological Research, vol. 6, no. 2, pp. 143–150, 2019.
30. S. Silvia Priscila, S. Rajest, R. Regin, T. Shynu, and R. Steffi, “Classification of Satellite Photographs Utilizing the K-Nearest Neighbor Algorithm,” Central Asian Journal of Mathematical Theory and Computer Sciences, vol. 4, no. 6, pp. 53–71, 2023.
31. S. Suman Rajest, S. Silvia Priscila, R. Regin, T. Shynu, and R. Steffi, “Application of Machine Learning to the Process of Crop Selection Based on Land Dataset,” International Journal on Orange Technologies, vol. 5, no. 6, pp. 91–112, 2023.
32. T. K. Behera, D. Marupaka, L. Thamma Reddi, and P. Gouda, “Enhancing Customer Support Efficiency through Seamless Issue Management Integration: Issue Sync Integration System,” European Chemical Bulletin, vol. 12, no. 10, pp. 1157–1178.
33. V. Rupapara, S. S. Rajest, R. Rajan, R. Steffi, T. Shynu, and G. J. A. Christabel, “A dynamic perceptual detector module-related telemonitoring for the intertubes of health services,” in Artificial Intelligence for Smart Healthcare, Cham: Springer International Publishing, 2023, pp. 245–274.
34. Veena and Gowrishankar, “Processing of healthcare data to investigate the correlations and the anomalies,” in 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2020.

35. W. C. H. Hong, “Improving English as a foreign language learners’ writing using a minimal grammar approach of teaching dependent clauses: A case study of Macao secondary school students,” in Innovative Approaches in Teaching English Writing to Chinese Speakers, B. L. Reynolds & M. F. Teng, Eds. De Gruyter Mouton, pp. 67–90, 2021.
36. W. C. H. Hong, “Macao Secondary School EFL Teachers' Perspectives on Written Corrective Feedback: Rationales and Constraints,” Journal of Educational Technology and Innovation, vol. 4, no. 4, pp. 1–13, 2021.
37. W. C. H. Hong, “The Effect of Absence of Explicit Knowledge on ESL/EFL Stress-Placement Accuracy: A quasi-experiment,” Asian EFL Journal, vol. 20, no. 2, pp. 262–279, 2018.
38. W. C. H. Hong, “The impact of ChatGPT on foreign language teaching and learning: Opportunities in education and research,” Journal of Educational Technology and Innovation, vol. 5, no. 1, pp. 37–45, 2023.
39. W. C. Hong, C. Y. Chi, J. Liu, Y. F. Zhang, V. N.-L. Lei, and X. S. Xu, “The influence of social education level on cybersecurity awareness and behaviour: A comparative study of university students and working graduates,” Education and Information Technologies, vol. 28, no. 1, pp. 439–470, 2022.
40. X. Lian, W. C. Hong, X. Xu, K.-Z. Kimberly, and Z. Wang, “The influence of picture book design on visual attention of children with autism: A pilot study,” International Journal of Developmental Disabilities, pp. 1–11, 2022.
41. X. Xu, W. C. Hong, Y. Zhang, H. Jiang, and J. Liu, “Learning paths design in personal learning environments: The impact on postgraduates’ cognitive achievements and satisfaction,” Innovations in Education and Teaching International, pp. 1–16, 2023.