**ChatGPT Friend or Foe?**

Rupesh Thapa

Abstracts

When was the paper published?

What was solved/happened?

What were the performance issues/measures?

What method did they adopt?

Are any issues left to be addressed?

Keywords: Chatbot, Chatgpt, Artificial Intelligence,

1. Introduction

Artificial Intelligence (AI) has revolutionized our everyday lives with brilliant operators that can perform different capacities. One such operator is the chatbot, an AI program that employs Natural Language Processing (NLP) and assumption investigation to reenact human discussion. Separated from entertaining individuals, chatbots are valuable in instruction, trade, well-being, and amusement. They can diminish benefit costs and handle numerous clients at the same time, giving locks in and effective help.

In trade, chatbots have ended up the favored channel for client administrations as they offer more locks in answers than inactive FAQ records. They can increase representative efficiency and energize more individuals to utilize administrations. Chatbots give comfortable and proficient help to clients, specifically reacting to their issues. Hence, chatbots are a basic innovation for businesses and people alike.

https://www.sciencedirect.com/science/article/pii/S2666827020300062#sec2

* 1. What is a chatbot? And short history

Chatbot, short for chatterbot, is a computer program that employs artificial intelligence (AI) and natural language processing (NLP)to reenact discussions with human clients through voice commands or content chats, particularly over the web. Chatbots are not as they were for amusement but are valuable in instruction, commerce, and e-commerce. They are well known due to points of interest for clients and designers, such as stage freedom, moment accessibility, solid installment integration, and notice frameworks. Chatbots can be coordinated into gather discussions, have constrained information prerequisites, and information is effectively transferable. Designers take advantage from communication unwavering quality, quick advancement iterations, and restricted plan endeavors.

Chatbot first ELIZA (named after the fictional Eliza Doolittle) was developed by computer scientist and MIT professor name Joseph Weizenbaum in the mid-1960s. As per the study, Alan Turing's paper on Computing Apparatus and Insights in 1950, presented the concept of the Turing Test to decide in case a machine had human-like insights. The entry at that point goes on to portray the improvement of a few chatbots, counting ELIZA, Repel, Jabberwacky, Dr. Sbaitso, ALICE, and SmarterChild, and their different applications and functionalities. The section moreover notes that Siri, a virtual right hand, was discharged by Apple in 2011 and employs common dialect handling and machine learning to perform different assignments.

https://www.sciencedirect.com/science/article/pii/S2666827020300062#sec2

https://link.springer.com/chapter/10.1007/978-3-030-49186-4\_31#Sec2

* 1. What are the significances of chatbot?

Chatbots are a pivotal mechanical progression (like Apple Siri and Alexa Amazon) that empowers businesses to communicate with clients utilizing fake insights and machine learning. As the number of chatbots on Facebook Messenger has expanded from 100,000 to over 400,000, it is obvious that this innovation is the favored channel for client administrations. By contributing in chatbots, companies can boost worker efficiency and empower more individuals to utilize their administrations. Chatbot helps in different sectors, for instance Improved Customer Service, Natural Language Conversations, Challenges and Opportunities, Systematic Literature Review. There are numerous significances of chatbot some are:

* Chatbots are the following vital mechanical headway that permits businesses to empower clients to communicate with informing frameworks with fake insights and machine learning innovation. This innovation is the foremost favored channel for client administrations.
* Chatbots can increase worker efficiency and empower more individuals to utilize administration.
* Chatbots have ended up progressively well-known due to the quick headway of the versatile gadget showcase and the advancement of informing stages over the past few long time, together with the included client benefits of having a recognizable interface, no requirement for downloading or introducing any additional application, and 24/7 accessibility.
* Researchers and developers have developed chatbots with different design techniques in response to advances in artificial intelligence, machine learning, and natural language processing techniques. This enables them to be more effective than conventionally designed chatbots.
* The ability of chatbots to understand user requests, process them, generate appropriate responses, and maintain a conversation with users remains a challenge, despite improvements in technology.

https://www.sciencedirect.com/science/article/abs/pii/S0957417421008745

* 1. What are types of chatbot?

Types of chatbot can be divided into three parts with its sub-group according to this study. Types of chatbot with its sub-group and its tasks are mentioned below in the table:

|  |  |  |
| --- | --- | --- |
| Types | Sub-Group | Tasks |
| Structure | Flow Chatbot  Artificial Intelligence  Hybrid | Flow chatbot means a tree-based chatbot which gives a fixed reply and only answers to questions which are already in the database set by developer. Flow chatbot incorporate button, catchphrases, and catchphrases rather than free composing to drive the client down the predefined way.  Chatbot with artificial intelligence has the capacity to overhaul its information and  recognition from past discussions and users’ involvement, letting the  clients lock in more openly.  This sort of chatbot combines the concepts of Flow and AI chatbots. This chatbot can get it and communicate with clients but remains within the  design decided by the engineer. |
| Purpose | Functionality  Fun | This chatbots have fixed capacities depending on the engineer (i.e., chatbot for  learning, individual right hand, update, online shop right hand, etc.).  chatbot that aiming as it were for amusement (i.e., recreations, funbot, etc.). |
| Audience | Generalist  Specialist | This chatbot has common information that we can inquire about specifically. I.e., Siri  created by Apple, and Cortana created by Microsoft. Both Chatbots can help us unravel common issues such as looking for eateries, areas, and more.  This chatbot centers on one compelled thing and do that one thing amazingly.  well (i.e., chatbots that are utilized to serve clients online when requesting things). |

https://www.sciencedirect.com/science/article/abs/pii/S0957417421008745

* 1. What is ChatGPT ? Discuss along with its history and differences with normal chatbots.

Chat Generative Pre-Trained Transformer (ChatGpt) is one of the powerful AI tool which is in high demand today’s date. ChatGpt is a machine learning algorithm demonstration created by Open AI, which has taken the entire world of Natural Language Processing (NLP) by storm to associate with clients. ChatGpt assists users by writing essays, e-mails, code, letter and so on also answer questions.

Firstly, there was a GPT-1 model discharged by OpenAI before the discharge of GPT-2. In 2018, GPT-1 was discharged with 117 million parameters, which was a noteworthy change over the dialect models that were accessible at the time. It was an essential milestone for the development of large-scale language models.

In February 2019, GPT-2 was released and was a noteworthy advancement over its forerunner, GPT-1. It contained 1.5 billion parameters, i.e., 10 times bigger than GPT-1. This bigger demonstrate estimate permitted GPT-2 to produce more coherent and relevantly important content.

GPT-3 was discharged in June 2020 and was indeed more considerable advancement over GPT-2. GPT-3 had 175 billion parameters, making it the biggest dialect demonstration ever made at the time of its discharge. GPT-3 was able to perform a wide run of common dialect errands, counting dialect interpretation, question-answering, and indeed composing imaginative fiction.

Due to huge success of GPT-3, GPT-4 was discharged in September 2021 and has indeed more parameters than GPT-3, with 6 trillion parameters. This expanded show estimate empowers me to create indeed more practical and coherent dialect, as well as perform more complex normal dialect handling errands.

ChatGpt apart from other chatbots is its ability to understand content. ChatGPT is its moved-forward exactness compared with conventional NLP apparatuses. Unlike conventional NLP models, which regularly depend on rule-based approaches and depend on human-defined lexicons and language structure, ChatGPT employments profound learning calculations to memorize from the information it is prepared on. This comes about in a demonstration that's competent of producing more human-like reactions, as well as recognizing designs in dialect that conventional NLP models may miss.

* 1. How chatGPT works? Explain with diagram.

ChatGPT could be an open device created by OpenAI that employs GPT innovation to perform a wide extent of text-based demands, such as creating common dialect reactions to basic and progressed questions, writing essays, and tending to efficiency issues. GPT could be a dialect show created by OpenAI that employs generative, unsupervised pretraining and discriminative, directed fine-tuning to refine its concepts. Its capacity to perform a wide extent of language-based assignments, counting interpretation, address replying, and content era, sets it separated. ChatGPT's normal dialect handling capabilities make it a perfect apparatus for dealing with essential client benefit requests, possibly valuable within the legitimate calling and might help with evaluating and giving input on understudy assignments. It reached over one million special clients within one week of its dispatch, demonstrating its value and ubiquity. The innovation may have critical impacts on different businesses and possibly diminish the time required to compose inquire about papers. Its versatility, capacity to utilize information effectively, and create human-like dialect make it an interesting and noteworthy development within the field of common dialect preparing and fake insights. This is how ChatGPT works:

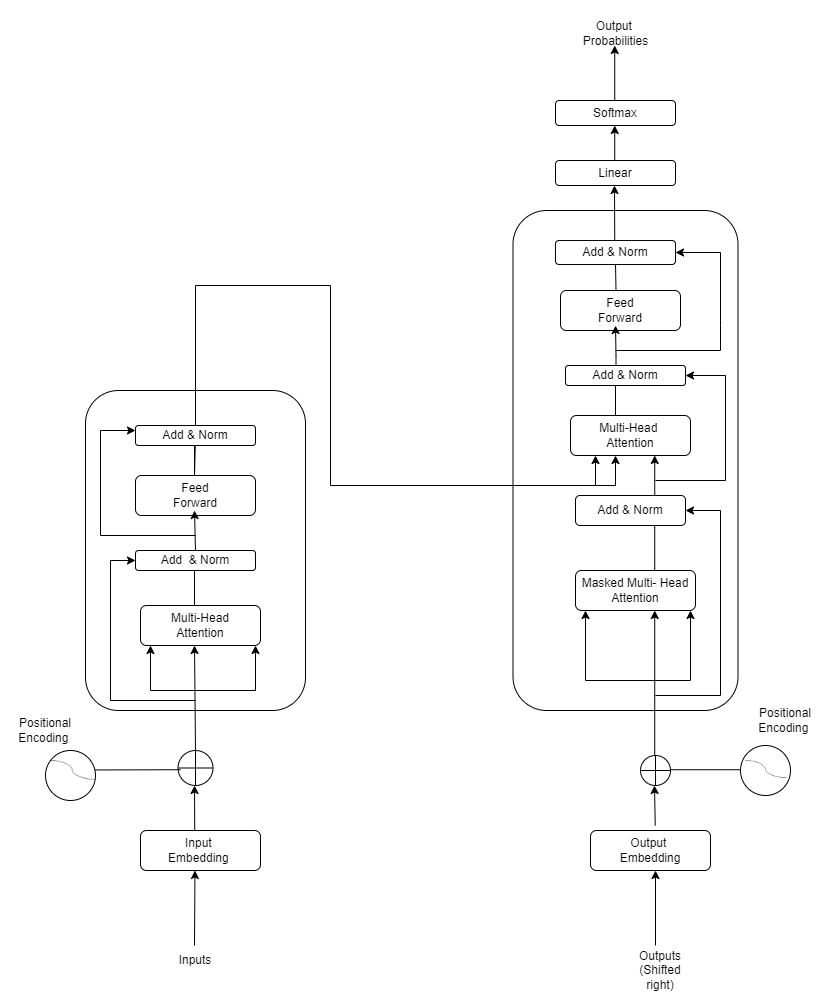


Figure 1: The Transformer – model architecture

Encoder and Decoder Stacks:

Encoder:

The design of the encoder in a transformer show comprises of N=6 indistinguishable layers, each with two sub-layers: multi-head self-attention and completely associated feed-forward. Leftover associations and layer normalization are utilized to assist stabilize preparation. All sub-layers and inserting layers create yields of measurement dmodel = 512 to empower the utilize of remaining associations.

Decoder:

The decoder within the transformer demonstrates N=6 layers, each with three sub-layers. It employs leftover associations, layer normalization, and adjusted self-attention sub-layers to go to known yields from the encoder stack. This guarantees forecasts for position *i* depend as it were on known yields at positions less than *i*.

Attention:

A compatibility function of the query with the corresponding key computes the weight attached to each value in the output.

Diagram

Description automatically generated

Figure 2: (Left) Scaled Dot-Product Attention. (Right) Multi-Head Attention consists of several attention layers running in parallel.

Scaled Dot-Product Attention:

Single dot-product attention is a type of attention mechanism used in neural networks, particularly in the Transformer architecture, which is commonly used for natural language processing tasks.

In single dot-product attention, there are three inputs: a query vector *Q*, a set of key vectors K, and a set of value vectors *V*. These three inputs are typically obtained from the output of previous layers in the neural network.

The attention mechanism computes the dot product between the query vector *Q* and each key vector *K*, and then applies a softmax function to obtain a set of weights that determine how much attention should be given to each value vector V. Specifically, the softmax function is applied to the dot product of *Q* and K divided by the square root of the dimensionality of the key vectors. This produces a set of weights that sum to 1 and can be interpreted as the importance of each value vector *V* for the given query *Q*.

The final output of the attention mechanism is a weighted sum of the value vectors *V*, where the weights are determined by the softmax of the dot product of the query vector *Q* and the key vectors *K*. This output is then passed to the next layer of the neural network for further processing.

Multi-Head Attention:

Multi-head attention is a technique utilized in neural systems, especially within the Transformer architecture, which is commonly utilized for natural language processing assignments. It expands the single dot-product attention mechanism to progress its expressiveness and capture more complex connections between inputs.

In multi-head attention, the input query vector Q, key vectors K, and value vectors V are to begin with linearly projected h times with distinctive learned linear projections to get h sets of questions, keys, and values, each with a decreased dimensionality. At that point, the dot product attention mechanism is connected to each of the h sets of anticipated questions, keys, and values in parallel, creating h sets of yield values.

The output values from each attention head are then concatenated and passed through another learned linear projection to get the ultimate output of the multi-head attention mechanism. This last output is at that point passed to another layer of the neural network for encouraging handling.

1. Literature review

The utilization of ChatGPT in education is still in its exploratory arrange and there's constrained investigation on its applications in this setting. However, a few significant articles have been recognized in Google Scholar, which recommends that ChatGPT can be an important asset in higher instruction. It can progress composing aptitudes, summarizing data, and diagram thoughts, sparing time and making strides in the quality of work. Besides, it can identify language structure and style errors, making the composed substance more comprehensible (Atlas, 2023).

Kasneci et al. (2023) found that ChatGPT can aid understudies to create inquiries about abilities by giving them data and assets on a specific subject. It can recommend unfamiliar perspectives and present understudies to modern investigation points, empowering them to pick up distant a much better; higher; stronger; improved" stronger understanding and assessment of the subject matter. Within the medical instruction field, Kung et al. (2023) have found that ChatGPT can help with clinical decision-making because it produces precise answers in restorative permitting exams. Rudolph et al. (2023) have distinguished a few focal points of ChatGPT, counting its capacity to produce human-like discussions, its speed and proficiency, and its cost-effectiveness since no human labor is required.

However, the utilization of ChatGPT in instruction has moreover raised a few concerns. As with any unused innovation, particularly when the assessment of information or aptitudes is intervened by innovation, questions have been raised almost the legitimacy of the learning involvement (García-Peñalvo, 2023). There are concerns that understudies may copy and glue writings created by ChatGPT without fundamentally analyzing what has been highlighted or chosen from a source, without citing the first sources, and without recognizing the potential for literary theft. This issue makes ChatGPT-produced content unacceptable for scholarly composing (García-Peñalvo, 2023). Issues of copyright infringement location in write-ups produced by ChatGPT have been raised, also how to recognize between reality and fiction content produced (Chatterjee & Dethlefs, 2023; Khalil & Er, 2023). Teachers are progressively stressed that understudies may utilize ChatGPT to create their composed assignments because it has been illustrated to produce reports in a matter of seconds without being identified by plagiarization locators. In any case, Atlas (2023) has contended that it could be a myth that unveiling the utilization of GPT-3 (dialect show made by OpenAI) would be considered plagiarization, and he showed that plagiarization really alludes to displaying somebody else’s thoughts as you possess without giving legitimate credit to the source. In this manner, when utilizing GPT-3, creators or understudies ought to make it clear that the show was utilized and cite or reference it fittingly.

Khalil and Er (2023) conducted an exploration to decide whether copyright infringement location instruments might identify expositions composed utilizing ChatGPT and found that of the 50 expositions tried, 40 had a similitude score of 20% or less, illustrating a tall degree of creativity. So also, Susnjak (2022) utilized ChatGPT in a test to evaluate its capacity to lock in basic considering instead of basically data recovery, and the comes about were exceedingly exact and exact, as well as consistently coherent. In differentiation, Dowling and Lucey(2023) famous that in spite of the fact that ChatGPT has points of interest for creating thoughts and recognizing information, it is weaker when it comes to writing blends and making suitable testing systems within the setting of fund inquiries about.

As a result of these concerns, a few schools have chosen to piece ChatGPT, as understudies may utilize it to consequently deliver assignments or other coursework (Ropek, 2023). Be that as it may, endeavoring to anticipate or boycott its utilization will not

address the basic issue of understudies looking for ways to balk the learning preparation. Instep, it is important for teachers to supply clear rules on the fitting utilization of ChatGPT, emphasizing the significance of basic investigation and appropriate quotations of sources. Furthermore, as the investigation on the applications of ChatGPT in instruction proceeds to advance, it is vital to investigate its potential benefits and restrictions, whereas guaranteeing that it is utilized in a moral and capable way. Eventually, the utilization of ChatGPT in instruction ought to be seen as a tool to improve understudy learning and engagement, instead of as an easy route or substitution for the learning handle.

Another zone that ChatGPT has appeared to guarantee is dialect learning and education. Concurring to Liu et al. (2022), ChatGPT can be utilized as an apparatus to bolster dialect learners by generating target dialect sentences that can be utilized for honing and assessment. This could offer assistance to learners to move forward with their composing and talking abilities by giving them a demonstration of sentence structure, lexicon utilization, and language structure. Also, ChatGPT can be utilized to naturally create questions and tests for dialect learners, which can offer assistance to evaluate their understanding of the dialect and recognize ranges for enhancement.

At long last, ChatGPT has moreover been utilized within the field of client benefit and back. Agreeing with Lin et al. (2022), ChatGPT can be used to make robotized chatbots that can help clients with their requests and issues. These chatbots can be prepared with a vast amount of information and can give fast and precise reactions to client questions, which can offer assistance to progress client fulfillment and decrease the workload of client benefit agents. Additionally, ChatGPT can be prepared to get its common dialect input and give personalized suggestions to clients, based on their past intelligent and buy history.

The technical implementations of ChatGPT incorporate the utilization of a transformer-based neural network architecture, pre-training with unsupervised learning, fine-tuning assignments with directed learning, and meta-learning. The transformer-based architecture incorporates an arrangement of self-attention layers that permit the model to capture the conditions between distinctive parts of a content arrangement. Pre-training with unsupervised learning includes preparing the show on a huge corpus of content information to memorize the fundamental structure and designs of dialect. Amid fine-tuning, the demonstration is prepared on a littler labeled dataset for a particular assignment, which permits it to adjust to the new task with fair many illustrations. A meta-learning approach, where the show is prepared to quickly learn unused assignments with constrained information by learning how to memorize. These specialized usages are what empower ChatGPT to be an effective few-shot and one-shot learner and to generalize well to a wide run of natural language processing tasks.

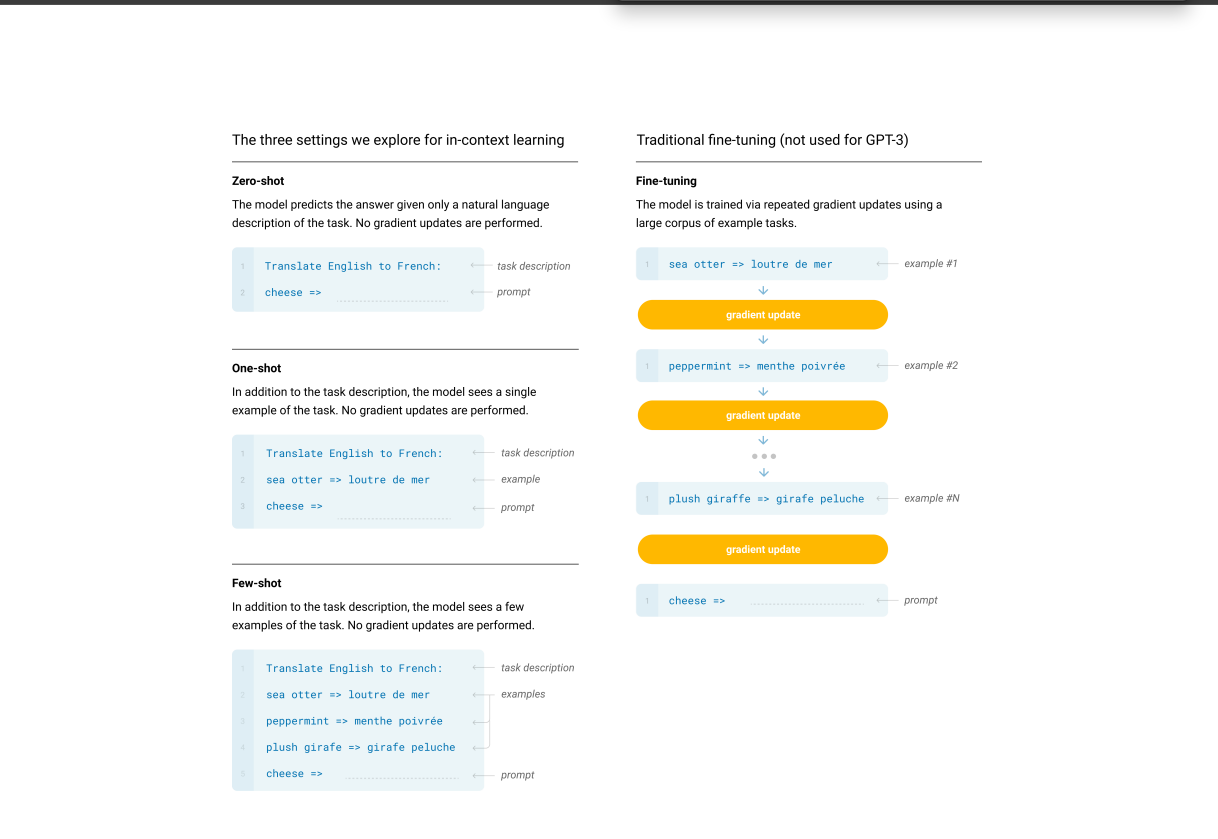


Figure 3: Context Learning of ChatGPT

A picture containing text, diagram, plan, technical drawing

Description automatically generated

Figure 4: Technical Implementation of ChatGPT

Fine-tuning (FT):

Fine-tuning could be a common approach where a pre-trained show is overhauled by preparing a particular dataset for a craved assignment. FT requires an expansive dataset and may result in destitute generalization and potential misuse of preparing information, driving an unjustifiable comparison with human execution. In this work, GPT-3 was not fine-tuned as the focus was on task-agnostic execution, but FT may be a promising heading for future work.

Few-Shot (FS):

Few-Shot alludes to the setting where the model is given many showings of the assignment at induction time without any weight overhauls. This strategy requires less task-specific information and diminishes the potential to memorize an excessively limited dispersion from a huge but contract fine-tuning dataset. However, the comes about of this strategy are much more regrettable than state-of-the-art fine-tuned models. A little sum of task-specific information is still required. This approach is related to few-shot learning in other settings in ML, where it involves learning based on a wide conveyance of errands and after that quickly adjusting to a modern assignment.

One-Shot (1S):

1s could be a setting whereas it where one demonstration of the errand is permitted, together with a normal dialect depiction of the errand. This setting is commonly utilized when assignments are communicated to people, such as when producing a dataset on a human specialist benefit like Mechanical Turk. The recognizing includes of 1S from few-shot and zero-shot is its closeness to the way errands are communicated to people.

Zero-Shot (0S):

0s includes as it were a common dialect instruction for an errand, with no exhibits permitted. This strategy is the foremost challenge but offers the greatest comfort and potential for vigor. Whereas it can be troublesome for people to get organized on an assignment without earlier cases, zero-shot is closest to how people perform a few assignments.

Zero-shot, one-shot and few-shot, contrasted with traditional fine-tuning:

This figure contrasts conventional fine-tuning with zero-shot, one-shot, and few-shot strategies for performing an assignment with a dialect demonstration. The last mentioned three strategies require as it were forward passes at test time, and within the few-shot setting, the demonstration is typically presented with a couple of dozen illustrations.

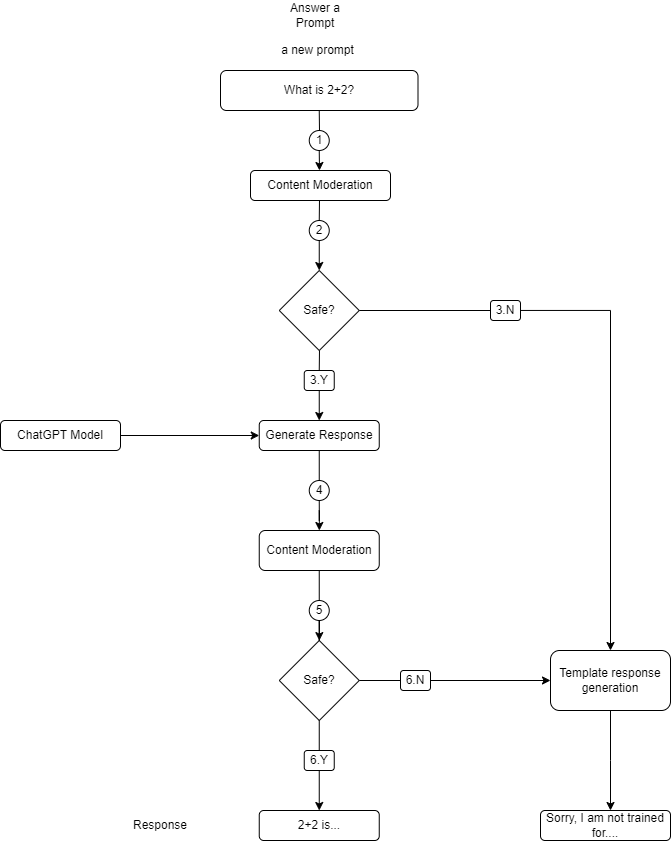


Figure 5: This figure shows, how ChatGPT gives the response?

ChatGPT has illustrated state-of-the-art execution in different investigations considers and real-world applications, counting dialect interpretation, content summarization, and so on. A few of the issues are:

Constrained spaces: GPT-3's preparing information is drawn from an assortment of sources, but it is still constrained to a certain set of domains. This implies that GPT-3 may battle with errands that require domain-specific information or language that's not well-represented in its preparing information.

Quality issues: Whereas GPT-3's preparing information is tremendous and assorted, it isn't safe to have quality issues. For case, the information may contain blunders or irregularities that can affect the model's execution.

Need for differences: Whereas GPT-3's preparing information is different in terms of dialect and points, it may still need differing qualities in terms of representation. For case, certain communities or societies may be underrepresented within the prepared information, which can lead to predispositions within the model's yields.

Need for control over information determination: GPT-3's preparing information is chosen naturally based on criteria such as significance and quality, but there's constrained control over which information points are included. This could make it challenging to address issues or inclinations within the preparation of information.

Restricted generalizability: Whereas GPT-3 can create amazing normal dialect yields, it may still battle with errands that require more generalizable information or thinking. For illustration, GPT-3 may battle with assignments that require common sense thinking or the capacity to draw associations between apparently irrelevant pieces of data.

Need for standardized measurements: There's no standardized set of measurements for assessing dialect models, which can make it troublesome to compare distinctive models. Measurements like perplexity and precision are frequently utilized, but they do not fundamentally capture all angles of a model's execution.

Restricted information accessibility: It can be troublesome to discover high-quality datasets that are expensive sufficient to supply a comprehensive assessment of a dialect demonstration. Numerous existing datasets are generally little or particular to a certain assignment or space.

// Difficulty of tasks: Some tasks, such as commonsense reasoning or natural language understanding, are inherently difficult and may not have clear benchmarks or gold standards for evaluation.

⎫ Model biases: Language models may exhibit biases that are not easily captured by standard evaluation metrics. For example, a model may generate offensive or harmful language even if its perplexity score is low.

⎫ Rapidly evolving field: The field of natural language processing is rapidly evolving, with new models and techniques being developed all the time. This can make it difficult to compare models that were trained using different methods or on different datasets.

Despite these challenges, there are efforts underway to develop more standardized benchmarks for language models, such as the SuperGLUE benchmark and the General Language Understanding Evaluation (GLUE) benchmark. By continuing to improve these benchmarks and evaluate models on a variety of tasks, we can gain a better understanding of the strengths and weaknesses of different language models and drive progress in the field of natural language processing.

Reivew of more than 10 papers from journal (2019-2023)

Discuss

ChatGPT and its applications

Industry applications>> education, health, medicine, industry, research

NLP Applications >> content generation, text summarization, machine translation, QA,…

chatGPT and its Pros / Cons / limitations

ChatGPT and its technical implementation with diagram

ChatGPT and its comparison with other AI-based chatbots

ChatGPT and performance issues

1. Analysis

Analyze using 10 or more papers

Critically analyze >> cross-check statement presented in one paper with that in another paper, present your opinion as well.

3.1 Ethical issues

3.2 Trust issues

3.3 Accountability issue

1. Conclusion

Present your final verdicts / recommendations based on your study