**Abstract**

According to recent studies, ChatGPT has the potential to be used for a variety of text annotation and text classification applications. However, identical input can produce many different outputs with ChatGPT because it is not deterministic, like human programmers. Considering this, it makes sense to judge the reliability of ChatGPT. To evaluate the consistency of ChatGPT's ability to classify and annotate text, this study focused on multiple model parameters, rapid changes, and repeated entries of the same data.

The results show that when used for a real-world classification task to divide Internet content into news and non-news, the consistency of ChatGPT output may not meet scientific standards. about reliability. For example, repeating the same entry or changing the wording of the prompt slightly can yield different results. Although accumulated results from multiple replicates may increase reliability, this study advises against using ChatGPT for zeroshot text annotations and emphasizes the need for widespread validation, such as comparisons with data indicating whether they were annotated by humans. ChatGPT should not be used for unattended text annotation and classification.

**Introduction**

As innovation creates and gets to be progressively coordinates into our lives, individuals commonly ponder in case AI frameworks will eventually be a companion or an foe. Chatbot, a sizeable dialect shows prepared by OpenAI utilizing the GPT-3.5 engineering, serves as an illustration of this.

Large languages (LLM) like BERT, GPT-3, and Codex have transformed a wide range of applications because of the fast advancements in natural language processing (NLP). These models have demonstrated superior performance in a variety of tasks, including text generation, code synthesis, and machine translation, among others. The most recent contribution to this model collection is OpenAI Chatbot, a generalized text model that has been pre-trained and enhanced with human input. Compared to preceding models that mostly depended on a single prompt, Chatbot offers especially strong conversational interaction capabilities, combining text generation with code synthesis. This study looks at if and how the capabilities of Chatbot may be used to the robotics industry.

**Significant aspects of Chatbot**

Chatbot is to demonstrate and test the capabilities of a powerful application system. Chatbot is a general AI program that uses language processing to create text, illustrations, music, and videos. Big language model supports Chatbot, however, it needs data to operate and evolve over time. As a person, a model learns to train he received. Algorithms become more adept at seeing patterns for predict future events and generate reliable text. He used a Sequence templates designed for text production tasks, including questions-and-responses, text summaries, and automatic translations. Chatbot can provide suggestions for goods or information that meet a specific need and preferences by studying customer data. The company can grow experience for new audiences, drive engagement, and build trust without Chatbot. For businesses looking to expand their customer base establishment, reach new markets, conduct effective marketing campaigns and forge closer relationship with existing and potential customers, ChatGPT can become a valuable tool. The important role of ChatGPT in Current scenarios are discussed in Table 1. ChatGPT can determine what makes a business successful inExploiting marketing tactics, customer base, product attributes, and other factorselement. It gives recommendations on how our company can applyor improve these characteristics. Looking at market trends, thecustomer expectations and other variables relevant to our business, ChatGPTcan suggest ways to take advantage of these opportunities and expand company by reviewing product offerings, marketing initiatives,and strategies to attract customers of competitors. goal-baseddemographics, marketing goals and budgets, ChatGPT can advise youon the best channels for a particular campaign. Social networkwebsites, email marketing, search engine marketing and othersMarketing channels are examples of channels. Performance analysishelp organizations monitor and evaluate the effectiveness of digital marketing activities by revealing what works and what doesn't.To improve results, campaign plans and tactics may be modifiedin real time thanks to this information

Major roles of ChatGPT

|  |  |
| --- | --- |
| Roles | Description |
| Learning and improving | * One of the main advantages of ChatGPT is its ability to learn from user interactions. * ChatGPT is a valuable tool for the creation and improvement of conversational AI systems in the future because to its adaptability to a variety of application situations. * A wide range of text-based materials, including books, news articles, websites, and more, were used to teach ChatGPT, giving it a deep grasp of a number of topics. |
| generating a lot of interest | * ChatGPT is a powerful tool that can dramatically increase human productivity and creativity, whether used to answer questions, inspire creative writing, or help with everyday tasks. * With its multiplex natural language processing capabilities, it has changed the way people interact with AI. |
| Business applications | * Businesses can develop more targeted marketing campaigns, engage with target audiences, and achieve their marketing goals when marketers use ChatGPT's capabilities. * To ensure that private, sensitive and confidential personal and business information stays where it is, use new technologies with caution. |
| Genuine conversations | * The responses from ChatGPT resemble those from actual talks in both appearance and tone. * The foundation of ChatGPT technology uses supervised and unsupervised AI learning methods to train some of the most well-known language models in the world. |
|  |  |

**How Chatbot works?**

ChatGPT is a kind of language model developed by Open AI. It uses a deep neural network architecture called Transformer for text generation. The model that takes a prompt or question as input, is preprocessed, and converted into a numeric representation known as a token. The input goes through a grid of encoders, computing a series of hidden representations or "encoders" to capture the meaning of the input. The decoder network will then use these encodings to generate the output text, one token at a time. The decoder is trained to predict the next token based on the input and the previous token generated by the model. The decoder network also has an attention mechanism, allowing it to focus on different parts of the input encoding when generating each token. This allows the model to selectively use information from the input to generate the output text. The output of the model is a string of tokens, which can be converted back to text and presented as a response. This model is trained on a large amount of textual data, allowing the model to learn language patterns and generate consistent and contextually relevant responses to different inputs.

**Graphical user interface

Description automatically generated with medium confidence**

**Model:** How ChatGPT works

After pre-training, the show was optimized for a specific movement, such as creating content or responding to request. To induce the show to work effectively on the specific work, the parameters ought to be fine-tuned. The demonstrate may learn the nuances of the assignment through this fine-tuning handle, which is able upgrade its execution.

Diagram

Description automatically generated

Figure 1: A specialist engineer must be included in the process to enhance current robotic pipelines. With ChatGPT, we intend to make it possible for (perhaps non-technical) people to interact with the language model via high-level language commands and successfully implement many platforms and tasks. any occurrence.

 The execution of the demonstrate was altogether progressed amid ChatGPT's advancement through broad testing and tuning. To make strides the model's precision and speed, OpenAI analysts tested with different designs, preparing strategies, and hyperparameters. In arrange to reinforce the model's capacity for creating coherent and changed answers, they too examined other preparing methodologies, counting the utilize of unsupervised learning approaches.

Diagram

Description automatically generated with medium confidence

The demonstrate was inevitably made accessible to designers by means of an API and discharged as a commercial item. As a result, software engineers were able to consolidate the model's normal dialect preparing capacities into their possess applications, such chatbots and virtual collaborators.

Discuss

**Trustworthiness**

Deep learning and large-scale models have profoundly changed AI in recent years. In areas like language and vision, AI models still have room for improvement. The cost and durability of large AI models, as well as privacy and trustworthiness point, pose other challenges. This course will explore the ethical advancement of deep learning and AI at scale in terms of language, vision, and multimodality, and examine its possible impacts on humans and society.

ChatGPT's primary objective is to help and take portion in human conversations. It is planning to supply data, respond to request, and take an interest in discourses on numerous subjects. In spite of the fact that a few would ponder whether ChatGPT is eventually a companion or enemy.

You can refer to ChatGPT as a buddy. It is a tool that may aid users in navigating challenging material, offer provocative question answers, and in some circumstances even serve as a buddy. For people, companies, and organizations wishing to enhance customer service or simplify processes, ChatGPT might be a useful tool.

On the other side, there are reservations regarding the possible dangers connected to AI systems like ChatGPT. Some people worry that as AI develops, it may eventually replace humans in a variety of professions, resulting in a significant loss of jobs. There are further worries regarding the moral ramifications of employing these technologies, in addition to privacy issues and biases in the data used to train AI systems.

In spite of these stresses, it's basic to be beyond any doubt that AI innovations like ChatGPT were eventually created by individuals for the advantage of people. As a result, depending on how they are utilized, they may either be partners or foes. We will make beyond any doubt that AI frameworks are inevitably a constrain for great in our society by embracing a cautious and moral approach to their improvement and utilization.

ChatGPT is a huge language show created by OpenAI utilizing the GPT (Generative Pre-trained Transformer) design. The improvement of ChatGPT included a significant sum of investigate and experimentation within the field of natural language processing (NLP).

 Pre-training a huge neural arrange on a sizable dataset of online content was the primary step within the creation of ChatGPT. The neural arrange was prepared to expect the taking after word in a given string of content amid this pre-training stage. The demonstrate may learn phonetic designs and relationships more successfully the more information and computing are utilized amid pre-training.

The Transformer engineering, which can oversee long-range connections and has been illustrated to be fruitful in NLP assignments, was received by OpenAI analysts. In arrange to empower the show to concentrate on different components of the input arrangement and capture complicated relationships and designs within the information, the GPT plan utilizes a number of self-attention layers.

Developers are now leveraging deep learning, machine learning, and natural language processing algorithms to create complex chatbots, as the creation of chatbots has advanced significantly.

One such advanced chatbot is ChatGPT (Chat Generative Pre-Trained Transformer).

Based on natural language processing, it was introduced by OpenAI in November 2022 and has since become extremely popular. Within four days, it had 1 million users, and two months after launch, it had 100 million. By comparison, TikTok took about nine months after its global launch to reach 100 million users, while Instagram took over two months to reach its first million signups.

 Technological developments are often a breakthrough, but the artificial intelligence debate has long divided supporters and opponents. Personalized shopping and learning, voice assistants, fraud detection, intelligent content creation, and driverless cars are just a few examples of applications for AI. The technology's main problems, despite its applications, are algorithmic bias, the (wrong) use of AI by authoritarian governments, and the existential threat posed by super-intelligent, self-improving AI. caused progress. We humans cannot control it. It is therefore not surprising that in its short existence, ChatGPT has caused excitement and fear among AI adversaries.

Panic spread naturally after ChatGPT was proposed in academia, including writing research grants, discussing new research directions, and writing research manuscripts. Publishers such as the American Association for the Advancement of Science (AAAS), which publishes the highly regarded journal Science, have now banned the citation of ChatGPT as an author and allowed the text to be published. copy. Copy. its version in scientific papers. Other well-known publishers also don't allow ChatGPT authoring in their work, but both of them allow the use of ChatGPT supposedly to improve the readability and language of articles. notebook. rescue .

In any case, it shows up that its application to and back for composing examine articles falls into a gray run. Though Springer Nature and Elsevier appear up to be fine with its utilize, AAAS has through and through denied it. wherein non-native English speakers might utilize manufactured insights (AI) fueled instruments like ChatGPT to upgrade the dialect and coherence in their inquire about articles. Dialect has long been a issue, or or maybe a obstruction, for logical distributions. ChatGPT may be able to level the playing field in terms of dialect in arrange to bolster the publication-based development of science.

The proliferation of "garbage science" may be a genuine issue, in spite of the plausibility that ChatGPT accidentally contributes to it. It may be a well-known reality that certain distributers are transparently unscrupulous and predatory8, exclusively concerned with their possess budgetary interface, and creators who distribute in these journals are as it were inquisitive about moving forward their distributing and quotation records. A combination like that's exceedingly terrifying, and ChatGPT has the control to significantly speed it up. Whereas the battle against predatory distributers and the rise of "junk science" will proceed, permitting ChatGPT utilization in composition composing with a clear say of this within the Affirmations area of the paper can be one potential arrangement for trustworthy distributers.

Another major problem we see with ChatGPT-assisted article preparation is citation and attribution of actual work referenced. The negative effects of secondary and tertiary citations, such as oversimplification and misinterpretation of the original work, have been identified by scientific research. Now it looks like ChatGPT can increase these appearances because, despite the fact that it can include citations of previously published work, the content it generates still needs Scientists work carefully to improve SEO. The worst case scenario is a huge list of fictional references that we sometimes see. Some examples of referenced manuscript introductions made by ChatGPT are included in the appendix

Diagram

Description automatically generated

**Abilities**

Take a closer look at ChatGPT's skillset for handling AI challenges in the field. Specifically, we evaluate ChatGPT's effectiveness in a variety of robotics-related tasks, from simple spatial-temporal thinking problems to practical deployments of aerial agents and manipulators.

We draw attention to a few notable talents we found in these tests.

Despite the immense capabilities of ChatGPT, the real security issues of the implementation should not be overlooked, especially when the implementation includes real bots. As Figure 2 shows, we think it's important to have someone in the loop to monitor things and take appropriate action in the event that ChatGPT generates unexpected behaviors.

**Ethics**

The roles of an ethics committee involve reviewing research proposals, ensuring that the study is conducted morally and legally, and offering assistance. This process comprises weighing the benefits and dangers of the research, getting informed permission, and monitoring continuing compliance.

The bureaucracy and lag time of ethics committees have been criticized, although these committees

They are not ethical, despite what their critics claim. Contrarily, the establishment of ethics committees is required to assure the protection of human subjects in research and to stop ethical infractions.

For the way ethics committees function, there are defined rules and principles. The Belmont report, for instance, emphasizes the moral values of justice, respect for people's rights, and other factors that go into doing research using human beings in a morally acceptable manner. Moreover, the shared feature

The Code includes rules for the protection of research subjects, such as the creation of ethics committees.

In recent years, there has been growing awareness of the necessity of improving accountability and openness in how ethics committees operate. This one sparked the creation of ethics committee criteria for moral behavior, including the Good Clinical Practice criteria and the Association for Accreditation Standard for Human Research Protection Programs (AAHRPP).

Experts from a range of disciplines, including representation from the domains of medical, law, ethics, and the community, make up the ethics committee. This multidisciplinary approach ensures that the examination of research proposals is thorough and takes into consideration the many stakeholder viewpoints.

Additionally, the development of policy depends on ethical committees. For instance, the National Institutes of Health (NIH) formed the Office for the Protection of Human Research (OHRP), which is in charge of developing and enforcing ethical standards for research relevant to civic learning.

In other words, doing research that is unethically conducted is not acceptable and requires the use of ethics committees. They must consistently seek to ensure accountability and openness in their activities in order to safeguard the rights and welfare of human subjects. Both scientists and the general public should be aware of the importance of ethics committees in promoting ethical use of scientific information and guaranteeing proper behavior throughout research.

**A STATEMENT OF ETHICS**

**GPT-3 Use That Is Ethical** The goal of this project was to avoid potential GPT-3 injury, and all of our GPT-3 tests were motivated by a desire to remember more and improve consistency. We believe our investigation will increase the stability of the model and make it safer to use. The ethical use of GPT-3 is mostly stressed in our segment of social disposition and politeness. We confirm that the calculation gives one-sided expectations, especially when the cases shown in the incentive program are statistically cosmetically unequal. Despite our efforts to respond to those trends, the show still has a long way to go in recent times, which makes perfect sense. Much more work is needed to achieve this.

**Limitations of This Work** We identified a number of gaps in the work and presented a list of unanswered questions for further investigation.

* **Other aspects of reliability:** We have focused on four key reliability characteristics of our work, while there may be other factors that we do not understand. examples include avoiding malicious and hallucinogenic builds, combating conflicts found by humans or red teams AI, and detecting harmful prompts like quick injection 5 (Ganguli et al., 2022 ; Branch et al., 2022 ; Perez et al., 2022) .
* **Techniques for enhancing trustworthiness:** Although we have completed the initial steps and found incentive techniques useful for certain reliability qualities, readers should not use our work as evidence that GPT -3 is now reliable and ready to deploy. Our tests really show that there is still room for improvement, such as reducing social biases and improving calibration. We believe our work will serve as a catalyst for more future research to develop more reliable LLM production methods.
* **Research to comprehend model behaviors:** Although the exact cause of these behaviors is unknown, we have discovered some remarkable properties of GPT-3. For example, what if the prompt, training data, training objective, or model design leads to a small generalization gap? What phrase or word causes GPT-3 to respond to prompts during depolarization? Why, even if providing anti-bias examples at the end of the prompt does not lead to recent bias (Zhao et al., 2021b), it can still lead to significant bias against minority groups? Can we infer behavior patterns from attention patterns or pre-training data? These studies can help us better understand how and why incentive functions, which will improve our LLM usability.