International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 12 Issue 2, February-2023, Impact Factor: 7.957

Language Model Analysis of GPT-3

Rohan Sarakinti

Research Scholars Program, Harvard Student Agencies, In collaboration with Learn with Leaders

ABSTRACT

Generative Pre-trained Transformer 3 (GPT-3) is a new language model introduced to the world recently by OpenAI, bringing with it many controversial opinions. GPT-3 has produced unbelievable results compared to other artificial intelligence models due to its vast capacity. GPT-3 is exceptionally versatile and specializes in text production, which many acknowledge. Alternate reviewers have claimed GPT-3 is overhyped and useless. This is mainly derived from GPT-3's ethical issues being heavily emphasized in today's society as awareness of the ethical aspects of AI is slowly increasing. Although GPT-3 has a lot of potential, the uncertainty of its results restrain it from being used in various fields. This exploratory research will touch upon the different aspects of GPT-3 and explain it on a surface level so no prior knowledge is needed. All of the testing mentioned in the paper was done by other individuals.

Keywords: GPT-3, language model, Artificial Intelligence,

Subject: Artificial Intelligence

INTRODUCTION

Technology has always been a vital part of our everyday lives allowing for communication, education, and even purchasing goods to become effortless. Technology has become a necessity for the functioning of life. As the world around us transforms, technology advances even quicker allowing us to do impressive things. One of the technologies empowering us to this level is Artificial Intelligence (AI). Generally, AI are programs that perform tasks that should require human perception and intelligence, such as judgment or decision making. Developers are now on a new journey to try and create an AI that perfectly replicates human nature and exhibits impeccable accuracy. Different AI programs have distinct features leading to the unique way in which they interpret data and give an output, so until we get the perfect AI program, we must understand the strengths and weaknesses of various programs at our disposal to judge if they are appropriate to be used in the real world and how exactly they should be utilized. Many are now analyzing a recent technological breakthrough in the AI space known as Generative Pre-trained Transformer 3. This study is centered around the different elements of OpenAI's GPT-3, but more specifically, this paper will delve into its exemplary features and limitations. Additionally, the potential applications of GPT-3 and the scope of its use will be addressed.

What is GPT-3?

Pre-trained deep learning models are fascinating in the way they take plenty of information and capture essential parts of it. After processing this information, they can then predict sequences and solve questions with contextual information (Zhang et al., 2021). Announced by OpenAI in June 2020, GPT-3 falls under the deep learning subset in the machine learning category, and is a 175 billion parameter autoregressive language model that constructs text as if written by humans (Floridi & Chiriatti, 2020). GPT-3 is a pre-trained program, which means it is trained with data sets of texts from a multitude of sources as more text makes it as relevant as possible. The sequence of data it generates is based on a source input, called a prompt (Floridi & Chiriatti, 2020). For example, if given a sentence as a prompt, GPT-3 would try to continue the sentences and create a paragraph related to the prompt by also utilizing the skills and knowledge gained when it was trained. GPT-3 is merely fitting general data that it learned and retained, and does not understand natural language yet, meaning it does not own the ability of logical reasoning (Zhang et al., 2021).

NEW AND IMPROVED

Large Scale

Compared to previously built programs, GPT-3 could perform on a much larger scale, strengthening its accuracy. GPT-2 had a negligible 1.5 billion parameter space and 40 GB data scale, whereas GPT-3 had 175 billion parameter space and 45 TB data scale. GPT-3 holds its position as the largest language model created as of early 2021, with WuDao 2.0 being the largest as of the time of writing this paper (Zhang et al., 2021). GPT-3's gigantic model capacity and immense training data tell us that it is a model supported by heaps of data leading to a better ability to generalize resulting in less dependency on supervised training (Zhang et al., 2021). When tested and compared against smaller models, it was found that with scale increasing, the zero-, one-, and few-shot performance grows, implying that larger scale models are more skilled meta-learners (Brown et al., 2020). This means that the language models gain a wider array of skills during training, including the ability to recognize trends or patterns. With the ability to accomplish downstream tasks without needing fine-tuning and multi-tasking general tasks, this program model only has room for improvement.

What Can it Do?

GPT-3 adapts and performs very well at various text-related tasks, such as answering questions, filling in missing text, translating words, and even doing 3-digit arithmetic (Brown et al., 2020). GPT-3 has also been used in other fields such as playing chess and writing web code which shows its versatility (Dale, 2020). Many companies and startups have recognized GPT-3's performance and are utilizing it to help build their company. Three common ways these companies are utilizing GPT-3 are for providing grammar services, autocompletion, and generating game dialogues.

Significant Flaws

Although GPT-3 has many powerful functions, the model still has some defects and ethical problems. As mentioned previously, GPT-3 does not have any logical reasoning, but rather depends on the quality and how well the data fits with the task at hand. One thing GPT-3 struggles to do well is to have semantic responses to a prompt, which is increasingly noticeable as the response grows longer and starts spewing out unrelated information (Floridi & Chiriatti, 2020). Consequently, when tested, GPT-3 was not up to par with fill-in-the-blanks, long text generation, and reasoning (Zhang et al., 2021). Another defect, that was tested by many others, is that GPT-3 can produce biased responses if trained using the wrong data, like low-quality internet data, and can even embody many stereotypes seen in various areas such as culture and gender roles (Floridi & Chiriatti, 2020). This further proves that GPT-3 only produces responses from statistical relevance with no regard to its meaning, leading to companies having to produce multiple responses and pick the one that conveys the message they want (Dale, 2020). Besides functionality problems, another concern the use of GPT-3 has is the high price needed to access more tokens, going up to $400 per month for 10M tokens, although there is a free version with 100,000 generated tokens (Dale, 2020). When a prompt is submitted and generated by GPT-3, it involves the process of making tokens. For the software to process the information, it converts the words of the prompt to tokens, and then converts the predicted tokens, or its reply, to words. Limiting the number of tokens means limiting the number of questions and responses you can give and receive from GPT-3.

Extent of GPT-3's Usage

Although GPT-3 is a truly developed tool made by professionals, its flaws limit its use in certain areas. GPT-3 is useful as an application in creative aspects, such as poetry and game dialogue, and it could prove to be very helpful in the mass production of low-quality pieces of text and might even be a helpful device for drafting these texts, such as recipes, summaries, and even newspaper articles (Floridi & Chiriatti, 2020). This can lead to many writers saving lots of time on drafts and spending more time on the editing process, leading to higher-quality final products. On the other hand, in the form of an expert system giving advice, GPT-3 is not suitable as it is still unreliable, and inaccurate responses in crucial fields such as the medical or even infrastructure industries could cause irreversible damage (Dale, 2020). Due to this, it is always better to have a human involved in the process when using GPT-3. Still, GPT-3 can considerably cut down time and might even help scour new innovative ideas found at the edges of the internet.

CONCLUSION & THE FUTURE

The role of AI in the current world is not yet solidly established, but GPT-3 promises to usher in a new era of technological development. Despite all of its shortcomings, GPT-3 is one of the best AI models available right now and is a major checkpoint in showing our advancement in technology. GPT-3's large capacity and ability to generalize data differentiate it from other programs. However, as new possibilities arrive and new doors are opened, this also allows for more issues to pop up. These issues, although important, can be minimized with proper human intervention. This tool should help the world see the amount of potential these evolved systems hold in our everyday lives, and allow

developers to build even more upon it. In the future, hopefully, innovators reach a point with AI where they have logical reasoning like never seen before. AI should reach a point beyond human intelligence where they transcend even beyond their data and training and help us find solutions to problems beyond what we could have ever imagined. To help attain this dream as quickly as we can, we need to start taking action and strive toward this better future, with the power of AI.

REFERENCES

[1]. Zhang, M., & Li, J. (2021). A commentary of GPT-3 in MIT Technology Review 2021. *Fundamental Research, 1*(6), 831–833. https://doi.org/10.1016/j.fmre.2021.11.011
[2]. Dale, R. (2020). GPT-3: What's it good for? *Natural Language Engineering, 27*(1), 113–118. https://doi.org/10.1017/s1351324920000601
[3]. Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines, 30*(4), 681–694. https://doi.org/10.1007/s11023-020-09548-1
[4]. Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in neural information processing systems,* 33, 1877-1901.