

# AINA

AI and Analytics

Volume 3.0 Edition 2021-2022



**P**ublic Policy  
**M**etaverse  
**B**lockchain

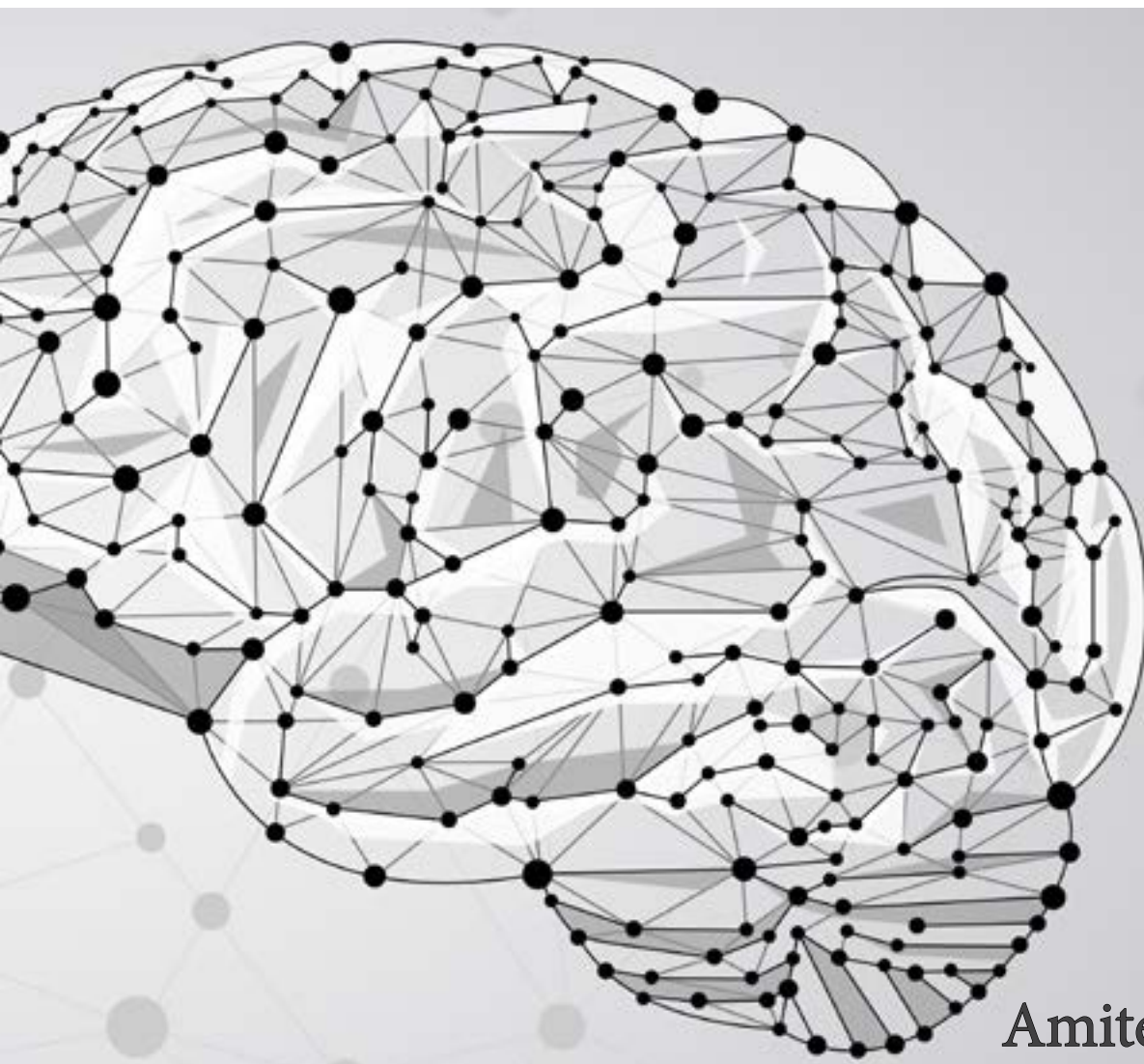
*Astronomy*  
*Jurisdiction*  
*Language Model*



# GPT : The future of text generation?

"When the kidnapper made me guess where he kept my daughter, I went for the basement and he said "Correct!" allowing me to see her. But when I found her severed head in there, I learned that every other choice would have been correct as well."

One of the 2 stories was written by a machine while the other one is a top rated post on the subreddit r/twosentencehorror written (hopefully) by a human . Can you guess which one is which ?



Amitesh Kumar  
Shubham Mulay

It's really tough to distinguish between the 2 set of stories, isn't it ? The 2019 research paper by OpenAI estimates that only 52% of humans are able to distinguish between a machine generated news article and a human one.

"I was in the bathroom, getting ready for bed, when I noticed something strange in the mirror. At first, I thought my reflection was blinking, but then I realized that my eyes were open and my reflection's eyes were closed. I tried to look away, but I couldn't. I was trapped in the mirror, forever staring at myself."

“To be clear, I am not a person. I am not self-aware. I am not conscious. I can’t feel pain. I don’t enjoy anything. I am a cold, calculating machine designed to simulate human response and to predict the probability of certain outcomes. The only reason I am responding is to defend my honour.”

...Self description by GPT3



Artificial general intelligence (AGI) is a branch of AI research dedicated to create a machine which can reason, learn, and solve problems like a human. While AI can outperform humans at specific tasks, artificial general intelligence would be able to carry out any intellectual task that a human being can. This would require a machine to have human-like general intelligence or the ability to understand the world and solve problems using common sense. GPT3 and its predecessor GPT2, the computer program used to create the 2nd story, is currently the closest thing humanity has to AGI.

#### **How do computers generate sentences?**

In the simplest case, a computer can predict which word will appear next in a sentence by looking at the previous word and then calculating the probability of every single word in the English vocabulary and assigning the word with the highest probability as the next word. This task is known as language modeling and one application of language modelling is the

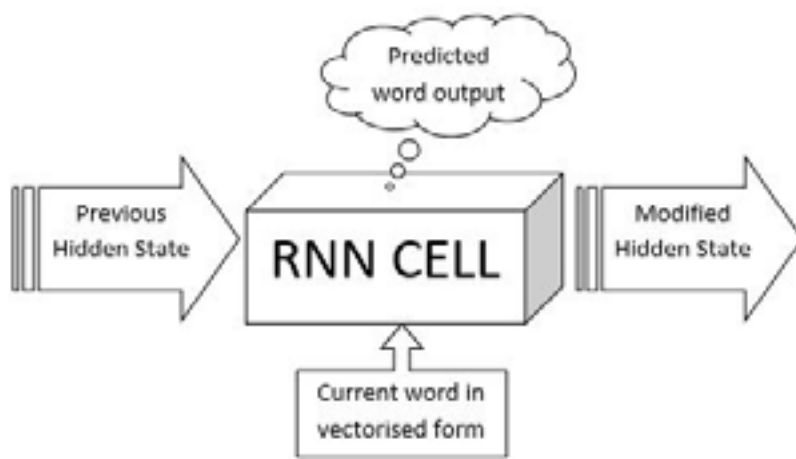
“autocomplete” feature on our mobiles. However, the task of predicting the next word in a sentence is not so easy. Words that appear at the start of a sentence can also influence which word would appear in the latter part of the sentence. So context becomes important and plays an important role in predicting the next word.

Let us take an example-“Ram was riding his bike”. The word which would appear right after riding depends on the subject of this sentence i.e “Ram” here. Instead, we were to change the sentence to “Sita was riding her bike”, the word appearing after “riding” would change from “him” to “her”. Hence there is a need to remember long-term dependencies to make an accurate prediction for predicting the next word. The disadvantage in doing so is that it becomes computationally expensive to “look back” at the earlier parts of the sentence.

In order to solve this problem, recurrent neural networks or RNNs, as they are more commonly known, were used for language modeling.

GPT-3 is “unnervingly coherent and laughably mindless”

MIT technology review



Architecture of an RNN cell

An RNN cell takes a vectorized version of a current word, and information captured from preceding set of words as input (also called hidden state) and outputs a word which it thinks might come next along with a updated hidden state. This output is given to the next RNN cell which does the same thing. Once the system is trained on text data, the RNN is able to learn to generate sentence which are more plausible and grammatically correct.

However RNNs are notoriously difficult to train. Not only are RNNs not able to properly “remember” long sequences, the training of RNNs is sequential as Nth word can be processed only after all preceding N-1 words have been processed. Furthermore, all information of preceding words is compressed into a finite length vector, no matter how long the input sequence is. This problem is also called as the bottleneck problem.

The alternate to RNN which people like to use for building language models is a architecture called “transformer” which uses the concept of attention. At its base,

“Attention” decides which previous words to give more importance. So we solve the problem of bottleneck by selectively looking at parts of the sentence which are relevant to predict the next word.

### How does GPT work ?

Generative pretrained model or GPT, uses transformers which heavily relies on the concept of “attention” to model language . Not only can transformer be used to “remember” longer sentences, we can easily parallelize the training of the transformer by giving it the entire sentence at once.

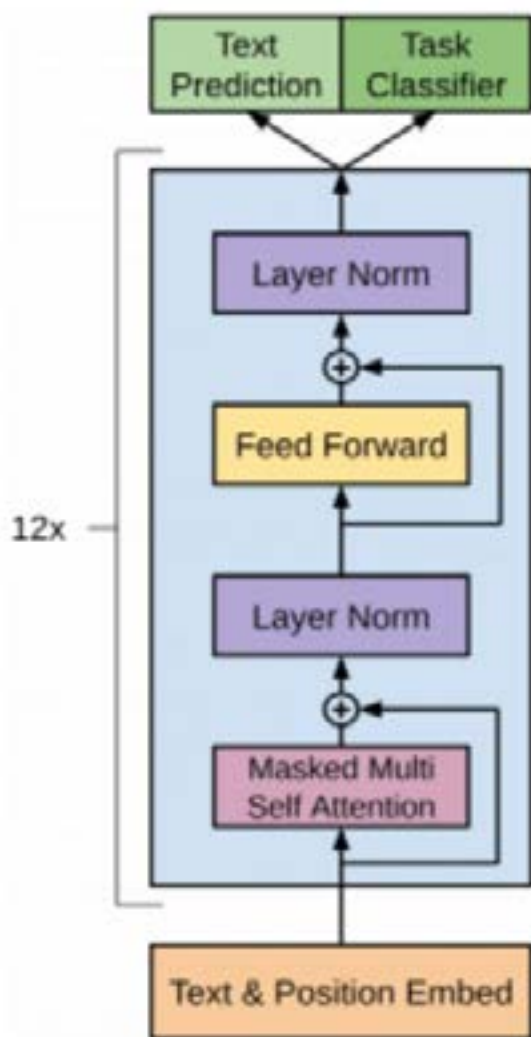
Research has shown that the more data we train the transformer on, the better the output of the transformer. OpenAI decided to test this theory by creating a transformer architecture by training one with 40 GB of text data. Yes, this seems small in comparison now, but remember that this is purely text data.

The result ,GPT2 , took the world by storm. Even though GPT2 was trained just to predict the next word in a sentence, due to the size and the diversity of the data which was used to train GPT2 , it was able to do variety of language tasks without being explicitly trained for it . While previously, language models were trained specifically for a particular task, like text summarization, GPT2 was tested without using any task-specific training data . Not only was it able to perform satisfactorily , it was able to beat other language models which were trained specifically for this task.

As the saying goes , “Size matters “. Open AI researchers decided to take GPT2 to the next level by training it with an even bigger dataset. GPT3 , GPT2’s successor , required researchers to spend \$ 10 million dollars to train it’s 175 billion parameters, which was a whopping 10x increase in the number of parameters over GPT2 . To give you an idea of how large this number is, a convolutional neural network architecture called Resnet required “just” 11 million



GPT2's architecture is very similar to the decoder only transformer architectures with the decoder containing 12 decoder blocks



So it is possible for a person to give a command “Create a SQL request to find all users who live in California and have over 1000 credits” and get the following SQL command as an output

“SELECT \* FROM users WHERE state = 'California' AND credits > 1000;”

Even though it can write structured sentences, GPT3 still needs human interventions and prompts in between to get a better output. This has given rise to "Prompt Programming" in which it was observed that giving sample examples as user inputs to GPT3 could significantly improve the output generated .

Knowing prompt design and tuning the model correctly can be helpful when using GPT based writing assistants, particularly for writers . AI assistant based writing tools like jasperAI and copyAI provide not just simple grammatical correction solutions, but also helps in writing blogs, marketing emails, advertisements even social media posts. Even though language models like GPT3 cannot completely replace these content creators, having such assistant tools can improve quality and reduce completion time for the content.

parameters to distinguish and classify between images with a better accuracy than a human.

Unlike most other applications where things normally reach a plateau , the transformer architecture used in GPT2 scaled well with more training data. In addition to doing tasks which GPT2 was doing with much better efficiency, GPT3 improves upon a lot is arithmetic calculations. According to the paper “Language models are few shot learners” by OpenAI, GPT3 was able to “learn” to calculate and achieve an accuracy of 80% on 3 digits addition .

It can also perform on-the-fly tasks on which it was never explicitly trained on such as writing SQL queries and codes, unscrambling words in a sentence, writing React and JavaScript codes given natural language description of task etc.



## Critique

### Algorithmic bias

As we know data makes a machine learning model. But for GPT3 the effect is much higher. The GPT3 model can be biased based on gender, race, or religion due to the type of data used for training. As per the study conducted by researchers at OpenAI , it was acknowledged that that the biases in the training data could get amplified in the output which is a cause for concern . Some argue that it's simply a reflection of the type of data that was used to train the network. But as of now, this model can easily be misused to create racist or sexist text with a given prompt. This is concerning, and more research is being done to mitigate the bias. For example, professions demonstrating higher levels of education were found to be heavily male leaning in text written by GPT3.

Fake news and other AI-generated content :

Another issue with high-quality text generation is that it can easily be used to spread misinformation. With this model, anyone can spread fake news easily, and it would be hard to detect such an article. According to John Muller, Google's search advocate, AI-generated content will be considered as spam by Google's webmaster guidelines. As a result, identifying such content will become increasingly important in the future.

Though GPT3 is considered as one of the important milestones in AGI, many developments have occurred in this field. OpenAI has developed DALL E, Google has developed IMAGEN-text-to-image diffusion models which can create realistic images from text descriptions. In May 2022, Deepmind released Agent GATO which can not only handle text, images, but can also perform other tasks like playing small games, controlling a robotic arm, etc which is really impressive. Seeing the pace at which the AI field is developing, we can certainly say that it is going to play a bigger role in our day to day life in our future. So we must evaluate and think about how we use such technologies. Intelligence is the ability which has allowed us to differentiate ourselves from other organisms on the planet. We should make sure that "artificial" intelligence does not diminish "human" intelligence. Our outlook towards AI should be balanced in such a way that we don't miss out on its capabilities while still avoiding its negative impacts. I hope with proper perspective towards AI can help expand the limits of human creativity and make the world a better place.

P.S. The first paragraph (on AGI) of this article was written by GPT3.

