ChatGPT/AI Generated text – Math Students’ Perspective

# AI Text Generation and ChatGPT – What is it, and how does it work?

Selina asked for a short write up about our experience with ChatGPT, but also maybe a brief explanation of how it works. Given that I’m a Computer Science student who is in the Machine Learning and Data Analytics specialization, I’m going to give a more detailed explanation first, then my perspective.

## ChatGPT – A brief history

Over the past couple months, ChatGPT has become the Kleenex or Google of AI generated text bots. It’s the go-to, name-brand implementation of the broader concept of an AI text generator. It was created by the company OpenAI, whose lab specializes in the research and development of AI tools. They currently produce several products, including DALL-E (AI image generation), Whisper (AI speech recognition), and GPT (*Generative Pre-trained Transformer*, used for AI generation of text). GPT was first proposed and created in 2018, based on a paper published by OpenAI’s Alec Radford.

ChatGPT is based on the newest version of this GPT model, GPT-3.5. All previous versions of GPT required the user to access it though another program, and had you interacted with it by sending specific, one-off requests. When developing ChatGPT, OpenAI altered certain parts of GPT-3.5 to allow for the conversational style that gives ChatGPT its broad appeal. It has a simple to use and easy to understand interface and allows you to converse with it in a very human manner.

## How does ChatGPT get its knowledge?

AI tools are always “trained” on a set of data, and then they can apply that training to other settings. GPT-3.5, the AI underlying ChatGPT, was trained on a very large set of data consisting of hundreds of billions of words. The data was made up of publicly available web texts (81%), books (16%), and all of Wikipedia (3%). In addition, it also knows how to write code in the most prominent languages. This gives ChatGPT a wide and varied array of knowledge, which is what gives it so much power. The biggest limitation, however, is that this dataset is from June of 2021. ChatGPT does not have any access to the internet, and as such can’t provide information about anything more recent than June 2021.

## Why has AI text generation become so prominent in recent months?

GPT has been around since 2018, and similar models have been around for at least as long. GPT-3.5, the underlying AI of ChatGPT, was first made publicly available in March of 2022, and can provide the same results as ChatGPT. So why was it only with the release of ChatGPT that AI generated text has become such a big deal?

The conversational style of ChatGPT has allowed it to become far more mainstream than anything previously has been. Although the same results could have been achieved for months before its release, ChatGPT has allowed more people than ever to access these tools. Students may have been using similar tools at Queen’s for years, but it just hasn’t been a prominent topic of conversation. ChatGPT does have some distinct advantages that make it more powerful, like the ability to converse back and forth with it to obtain better answers, but it is largely the same responses that you could get with GPT-3.5.

## Where do we go from here – what’s coming next?

In the coming months and years, we will see many similar AI chat bots pop up. Microsoft has licensed GPT-3.5 for use in its own bot, which is currently in a public testing phase. “Bing Chat” has one big advantage over ChatGPT – it has internet access. It has the entire knowledge base that ChatGPT has but can also refer users to external articles and even cite its sources. Bing Chat will likely become available to anyone in the coming weeks. Google is working on its own similar chatbot, called “Bern,” and there will surely be many more.

Another upcoming innovation will be GPT-4. It’s supposed to launch later this year, although there is no official announcement yet. Either way, the world of AI text generation and interaction is fast moving, and must be monitored closely from an academia perspective, to know what kinds of tools students may or may not be using.

# My Perspective as a Student (Specifically Computer Science):

## My Biggest take away:

Overall, the most important thing that I think many students need to know is that ChatGPT is always confident in its answers and will often be **confidently wrong.** It must be laboriously fact checked, as it will not give you any indication that it is unsure about something. It states everything as a fact.

### Written assignments:

Though I have never submitted any work from ChatGPT (I swear!), I spent a day running a variety of old assignments through it when the service first launched to get a feel for what it can and cannot do. There are certain assignments that it’s very good at writing, and others that it really struggles with. For written assignments, the results vary based on the topic you’re exploring and the amount of text you want it to generate.

If you ask for an explanation of something, it can almost always do a good job. For example, I had it generate a structured proof for the commutative property of a given relation in my Discrete Mathematics class. When you ask an open-ended question like “prove \_\_\_,” it can usually generate a fairly good answer. The answer it gave me was entirely correct and could’ve acted as a good blueprint for the assignment. If you ask it to explain a topic to you, it will almost always do a good job, even with obscure subjects. This can also guide a written assignment.

When asked to write an entire essay, my results have been mixed. For example, I kept using prompts similar to “Write a 1000-word essay about \_\_\_\_\_.” When I asked it for topics like “Why Puppies are loved by so many people” or “The history of the band Depeche Mode,” I almost always got reasonably good results. I did, however, run into big problems with more specific topics. My housemate, for example, asked it for a biology topic. The resulting essay was essentially 4 sentences repeated over and over to pad the essay to 1000 words, because ChatGPT simply didn’t know enough about the topic.

One of its best uses is asking for it to write an outline for something. It can almost always hit all the important points for you, and then you must manually fill in the blanks to create a well-structured essay. And of course, you can ask it to refine its output as well using follow-up questions.

## Programming Questions

Programming questions are one of ChatGPT’s strong suits. It will write fully working code that can simply be copied, pasted, and run for an assignment. It knows the most prominent languages, including Java and Python that my department uses in several courses. The problem with this “write a whole program” approach, however, is that I have no idea what’s going on in the code or how to change it if need be. It can take about as long to read and understand the code as it would’ve for me to just write the code myself in the first place.

The best use for ChatGPT with programming is to ask for help along the way. It can lay out in plain English what you need to do in the program, or it can tell you how to write a specific algorithm. I asked it to tell me how to implement Dijkstra's Algorithm, for example, and it told me exactly what my code would have to do without providing any actual code.

## Math Questions:

ChatGPT is mainly a text generation tool, and as such struggles with math questions sometimes. It can explain things to you conceptually but isn’t great at just solving numeric problems. Copying and pasting math problems into it is less useful than with other types of problems.

## Music Questions:

This was the biggest shock for me when first testing ChatGPT. I asked it to “Write an 800-word essay about the musical differences between Depeche Mode’s World in My Eyes and Claude Debussy’s Claire De Lune”. This was a very specific question about two pieces of music and given that ChatGPT can’t listen to audio I thought it was going to have a real hard time with it. The output, however, was quite comprehensive and accurate.

What I believe was going on is that it had access to articles written about both pieces of music, and so it could grab information about the musical aspects of both. While the essay was a little repetitive, it would’ve probably gotten a 75% or so. Overall, I was very impressed by the results.