

Data and the Human: Professor Fyfe Turing Test

Materials (with AI) – The tool that I am using to write this essay is called GPT-2. It gets its training data from the University of Chicago's The Dreaming Machine: Data from the Cognitive Science of Learning and Memory. I did not know about this tool at the start of the year because I was not sure what it was or how its strengths and weaknesses might be applied. GPT-2 gets much of its training data from Reddit, and compiles user's posts in order to form sentences and predict algorithms for what comes next. It uses some machine learning algorithms and first-order logic to meet the challenge of finding sentences that are most likely to be used in the future. This is the first time I have used this tool and it has been a very good learning experience. The tool is very easy to use and has a lot of documentation. It is designed as a quick tool that can be stored in a text file, or in a database on a computer, and then easily accessed from various web sites. I have found that once I have the data in my file, I can easily download it and use it to write my essay. However, I have found the tool too cumbersome to use at times, especially if I am just trying to write an essay. It is easy to use, but if you are trying to write an essay, you may find that it is not simple to navigate the tools and change frequently needed values.

Methods (with AI) – The way that GPT-2 works is by using machine learning and deep AI learning in order to generate more efficient solutions that can be used on an internal blog. First, the algorithm was initially designed to generate blog posts. However, once it had been developed enough it could also be used to generate 3D models, text and images. Like I said, it can be used for many purposes. Perhaps most importantly, it's embedded in a deep learning framework developed by Google called TensorFlow. This way, it's available to developers who want to use it to create their own applications. In addition to that, the GPT-2 is also a part of the Google Cloud Platform that can be used by any developer for creating their own applications. It's a great platform for creating everything from a simple app to a full-fledged online service. Some of the settings such as text length and prompt I played around with. I put different prompts into GPT-2 and set it to have different examples and picked the best, most human-sounding sentences that made sense. It turns out that GPT-2 has some very good features that are not in GPT-1. One thing that I noticed is that the commands that prompt you to type in a new row are not always the ones that you type in. In GPT-1 you can type in a new row and then hit enter to read the next row. In GPT-2, you have to hit enter multiple times to get the next row.

It was very interesting, GPT-2 did sound human-like, however it was obvious that when asked for multiple examples using the same prompt, the examples got more and more random and sounded a lot worse, not speech wise, but logic wise. It began to go on tangents that were slightly off topic, but made sense, so I had to narrow the word count and cut out parts of what the AI produced. I prompted it with whole sentences and even half sentences to see what it finished the sentence predicting. I then asked it to predict with a sequence of words, and it produced whole sentences that were completely unrelated to the sequence, and then it would predict a sequence of words and then produce a whole sentence that it knew what it was talking about, but it still would output random nonsense. I had to narrow the topic for it to stop generating random nonsense. When I noticed that it was able to predict a sentence like a yet-to-be-written movie script, it was like a light bulb went on. It knew the script and was able to predict the rest of the sentence, and then it was able to predict a sequence of words to the point that it was able to

predict many sentences. I was amazed, and I was so happy to have stumbled upon this. I was able to take an AI that was being trained on stories and stories that were written, and turn it into something that could produce results that were similar to my own writing.

Discussion (with AI) – When we first started this module, I remember talking about the Turing test, watching the movie “Her” and thinking to myself, there is no way AI could get to be that human-like anytime soon - that we do not have much to worry about. I was clearly wrong. The Turing test is not about choosing between organic and man made, it is about choosing the best algorithm that could bind both humanity and AI. Perhaps at the end of the day, AI will be no different, it will still just be a machine - the question then becomes, what happens when AI becomes better? If the AI becomes smarter than humans, then how do we deal with it? This is where we find the Turing test. The main idea behind this test is that the AI must have perfect intelligence to be able to do what humans do. If it is not perfect, then it is not intelligent enough to be able to write computers or call them as we know them today. If the AI has to be smarter than us, then we have to decide if we are willing to accept a machine with imperfect intelligence as it is. This test is incredibly important as it tells us whether or not we are willing to accept a machine that is smarter than us to have a part in our lives. Not that we will all be falling in love with AI in the near future, as seen in the movie, but the potential problems with privacy, AI bots becoming more and more tailored to humans and collecting information about them such as purchases made can be seen to rise. This could be harmful in the sense that the AI could then use algorithms that were man-made and control the market. I think a huge key take away from all of this is that AI is man made. Humans are biased, always, therefore AI is biased as well.

In the novel we read, “Speak”, it is shown many times that the AI is a collection of human voices and thoughts. Humans are also a collection of what we are told and perspectives we are fed, so we are biased, but that is known - I think a danger that comes with AI is that, to many people, these inherent biases are unknown and they may put trust into AI that could have skewed data or information fed to it. AI is going to be the world's most important technology. If we are to have an impact on the world, it is important that we understand what is happening with AI. We can't be experts on everything and everyone will have a different view of the future, but knowing the potential threats it can have on the world is very important. As seen in another one of the articles we read, “New AI Fake Text Generator May Be Too Dangerous to Release, Say Creators”, it is discussed that if GPT-2 got in the wrong hands, it could affect the economy and businesses by leaving negative or positive reviews, swaying other people’s bias, it could create even more fake news, as problem within society that is so prevalent right now. The article says, “OpenAI made one version of GPT2 with a few modest tweaks that can be used to generate infinite positive – or negative – reviews of products. Spam and fake news are two other obvious potential downsides, as is the AI’s unfiltered nature . As it is trained on the internet, it is not hard to encourage it to generate bigoted text, conspiracy theories and so on.” If the AI gets the hands of people with malicious intent, there could be so many problems escalating even more than they already are. Throughout this class, we have touched on so many topics and I think the main takeaway I got from this topic is that machine learning is more prevalent and already more human-like than I think - which has a lot of benefits but also a lot of drawbacks and potential concerns. After reading the article “Can a Machine Learn to Write for The New Yorker?,” it

became clear to me that AI is far more advanced than many people are aware of. This article highlighted how GPT-2 could make writing into something that is not necessarily the author's own and would inherently eliminate the original, authentic thought that goes into the act of writing - such as what I am partially doing right now with this paper. This article closes by saying, "Humans would stop writing, or at least publishing, because all the readers would be captivated by the machines. What then? GPT-2, prompted with that paragraph, predicted the next sentence: "In a way, the humans would be making progress." This was a great way to end the article and really depicts the controversy of all of this - what defines making progress by the human race? Being able to manipulate machines better? Doing things faster and more efficiently? Or is it keeping the authenticity of the human race while sacrificing efficiency? These are just a few of the questions that can align with the problems that society faces while trying to balance the growing use of AI.

Analysis (without AI) – Even though the AI wrote over half of my essay, it definitely took longer because I had to reword my prompt a lot and change my own words so that the AI knew enough of what I was talking about to be able to continue the sentence or continue speaking on the same topic for a while. I also had to read and sift through the different samples and pick which one sounded most human like or made any sort of sense, along with having to play with the word count a lot, which also accounted for this essay to be more difficult than actually writing it. I think it is impressive that with some guidance, the AI can produce a pretty human-sounding essay, but I think it will be more impressive once it gets more accurate. I do not think it will truly be able to write an essay any time soon as well as a human and have it make sense without a human checking for logic errors - a lot of the sentences sounded great but had cyclical thought processes and still lacked the intelligence of actually understanding what each sentence and word means. The AI understands how to connect patterns and use good grammar to continue a topic, but it cannot logically think through and understand the steps it takes to write a well thought out essay nor understand what the sentences that are being produced actually mean. I think GPT-2 worked really well, with my own guidance, but if I had set the AI to 1500 words to make it write this entire essay, it would have rambled down a weird path that would have made no sense - I toyed with the word count and above about 400 words it got pretty off topic. It definitely resembles my own casual style of writing, uses first person a lot, and I noticed that when I was playing with it and typed in slang such as "wuddup my name is abby" it changed the tone of the output and clearly got some slang from Reddit and it continued in all lowercase which was cool to see. I also noticed that GPT-2 had some grammatical errors, one of the statements in paragraphs where I used GPT-2, I had a problem with it using a plural version of "services" when it said "to a full-fledged online services" instead of "to a full fledged online service", which I thought was interesting to see that the AI still makes errors, although definitely way less than what I, being a human, would make. It was much easier to depict deep thought and logic within my essay without the AI - with the AI being used the writing seems much more shallow and redundant.

I would say that overall, I would honestly prefer to write a paper without the help of AI. Although it did not necessarily feel like cheating, it did leave a poor taste in my mouth because much of it was not my own authentic thought. As I talked about above, I spent more time toying

around with what I prompted the AI with to get it to say something that made sense and sorting through what was logical and what was not, than I would have actually just sitting down and writing the paper. Not that it wasn't faster than what my brain could produce, it was very fast in outputting many different examples of what to say, but it needed the guidance of a human and by the time all was said and done, I honestly think my brain would have been better - and this is coming from a college sophomore sitting in her bed writing this after having major COVID brain-fog. I would say this paper is not my writing - this analysis is - but the paper as a whole lacks my style and my authentic personality shown through my writing. I think I reader may recognize the difference in parts where I write vs GPT-2 because of the redundancy and because this paper requires lots of logical human thinking to put into it, but as GPT-2 gets more accurate and if the prompt was simply to write a personal narrative about myself, I think GPT- 2 would have done a better job than me. It might make up a whole different life and alter-ego, but if it came down to making up random scenarios after prompting it with "My name is Abby and I...", it would come up with crazy, awesome stories that were realistic enough that most people probably would not be able to tell the difference. I would probably use this tool again if I am stuck in a creative block for writing or want some good ideas with how to continue talking, or even if I just wanted to fluff a paper up for word count sake, but ultimately, I would definitely not use this to write my essay for me again and force myself to incorporate random things the AI spews out - I think I would trust my own brain more with that task. This experiment opened my eyes again to not just the potential AI has to develop in the future, but the broader question of what differentiates the human race and the machine - what part does having humanness and humanity play in getting our ideas across and how different is it when something is written by a machine? I do not think that right now, AI comes close to the humanness that a real human brain has, but maybe with time and if AI begins to eventually think for itself, it would sound more human - but of course, once AI becomes identical to human life, what is the need for the human race? And if the AI is able to think for itself and realizes this question of why have a human race if AI is the same thing, but more intelligent, more efficient, less room for error, and can start to build and train themselves? The next question may be: why not eliminate the human race as a whole? What is the true need for the human race? And not to end this essay on a bad note, but that is a very very scary thought.

Appendix –

Materials (with AI) – The tool that I am using to write this essay is called GPT-2. It gets its training data from the University of Chicago's The Dreaming Machine: Data from the Cognitive Science of Learning and Memory. I did not know about this tool at the start of the year because I was not sure what it was or how its strengths and weaknesses might be applied. GPT-2 gets much of its training data from Reddit, and compiles user's posts in order to form sentences and predict algorithms for what comes next. It uses some machine learning algorithms and first-order logic to meet the challenge of finding sentences that are most likely to be used in the future. This is the first time I have used this tool and it has been a very good learning experience. The tool is very easy to use and has a lot of documentation. It is designed as a quick tool that can be stored in a text file, or in a database on a computer, and then easily accessed from various web sites. I have found that once I have the data in my file, I can easily download it and use it to write my essay. However, I have found the tool too cumbersome to use at times, especially if I am just trying to write an essay. It is easy to use, but if you are trying to write an essay, you may find that it is not simple to navigate the tools and change frequently needed values.

Methods (with AI) – The way that GPT-2 works is by using machine learning and deep AI learning in order to generate more efficient solutions that can be used on an internal blog. First, the algorithm was initially designed to generate blog posts. However, once it had been developed enough it could also be used to generate 3D models, text and images. Like I said, it can be used for many purposes. Perhaps most importantly, it's embedded in a deep learning framework developed by Google called TensorFlow. This way, it's available to developers who want to use it to create their own applications. In addition to that, the GPT-2 is also a part of the Google Cloud Platform that can be used by any developer for creating their own applications. It's a great platform for creating everything from a simple app to a full-fledged online service. Some of the settings such as text length and prompt I played around with. I put different prompts into GPT-2 and set it to have different examples and picked the best, most human-sounding sentences that made sense. It turns out that GPT-2 has some very good features that are not in GPT-1. One thing that I noticed is that the commands that prompt you to type in a new row are not always the ones that you type in. In GPT-1 you can type in a new row and then hit enter to read the next row. In GPT-2, you have to hit enter multiple times to get the next row.

It was very interesting, GPT-2 did sound human-like, however it was obvious that when asked for multiple examples using the same prompt, the examples got more and more random and sounded a lot worse, not speech wise, but logic wise. It began to go on tangents that were slightly off topic, but made sense, so I had to narrow the word count and cut out parts of what the AI produced. I prompted it with whole sentences and even half sentences to see what it finished the sentence predicting. I then asked it to predict with a sequence of words, and it produced whole sentences that were completely unrelated to the sequence, and then it would predict a sequence of words and then produce a whole sentence that it knew what it was talking about, but it still would output random nonsense. I had to narrow the topic for it to stop generating random nonsense. When I noticed that it was able to predict a sentence like a yet-to-be-written movie script, it was like a light bulb went on. It knew the script and was able to predict the rest of the sentence, and then it was able to predict a sequence of words to the point that it was able to

predict many sentences. I was amazed, and I was so happy to have stumbled upon this. I was able to take an AI that was being trained on stories and stories that were written, and turn it into something that could produce results that were similar to my own writing.

Discussion (with AI) – When we first started this module, I remember talking about the Turing test, watching the movie “Her” and thinking to myself, there is no way AI could get to be that human-like anytime soon - that we do not have much to worry about. I was clearly wrong. The Turing test is not about choosing between organic and man made, it is about choosing the best algorithm that could bind both humanity and AI. Perhaps at the end of the day, AI will be no different, it will still just be a machine - the question then becomes, what happens when AI becomes better? If the AI becomes smarter than humans, then how do we deal with it? This is where we find the Turing test. The main idea behind this test is that the AI must have perfect intelligence to be able to do what humans do. If it is not perfect, then it is not intelligent enough to be able to write computers or call them as we know them today. If the AI has to be smarter than us, then we have to decide if we are willing to accept a machine with imperfect intelligence as it is. This test is incredibly important as it tells us whether or not we are willing to accept a machine that is smarter than us to have a part in our lives. Not that we will all be falling in love with AI in the near future, as seen in the movie, but the potential problems with privacy, AI bots becoming more and more tailored to humans and collecting information about them such as purchases made can be seen to rise. This could be harmful in the sense that the AI could then use algorithms that were man-made and control the market. I think a huge key take away from all of this is that AI is man made. Humans are biased, always, therefore AI is biased as well.

In the novel we read, “Speak”, it is shown many times that the AI is a collection of human voices and thoughts. Humans are also a collection of what we are told and perspectives we are fed, so we are biased, but that is known - I think a danger that comes with AI is that, to many people, these inherent biases are unknown and they may put trust into AI that could have skewed data or information fed to it. AI is going to be the world's most important technology. If we are to have an impact on the world, it is important that we understand what is happening with AI. We can't be experts on everything and everyone will have a different view of the future, but knowing the potential threats it can have on the world is very important. As seen in another one of the articles we read, “New AI Fake Text Generator May Be Too Dangerous to Release, Say Creators”, it is discussed that if GPT-2 got in the wrong hands, it could affect the economy and businesses by leaving negative or positive reviews, swaying other people’s bias, it could create even more fake news, as problem within society that is so prevalent right now. The article says, “OpenAI made one version of GPT2 with a few modest tweaks that can be used to generate infinite positive – or negative – reviews of products. Spam and fake news are two other obvious potential downsides, as is the AI’s unfiltered nature . As it is trained on the internet, it is not hard to encourage it to generate bigoted text, conspiracy theories and so on.” If the AI gets the hands of people with malicious intent, there could be so many problems escalating even more than they already are. Throughout this class, we have touched on so many topics and I think the main takeaway I got from this topic is that machine learning is more prevalent and already more human-like than I think - which has a lot of benefits but also a lot of drawbacks and potential concerns. After reading the article “Can a Machine Learn to Write for The New Yorker?,” it

became clear to me that AI is far more advanced than many people are aware of. This article highlighted how GPT-2 could make writing into something that is not necessarily the author's own and would inherently eliminate the original, authentic thought that goes into the act of writing - such as what I am partially doing right now with this paper. This article closes by saying, "Humans would stop writing, or at least publishing, because all the readers would be captivated by the machines. What then? GPT-2, prompted with that paragraph, predicted the next sentence: "In a way, the humans would be making progress." This was a great way to end the article and really depicts the controversy of all of this - what defines making progress by the human race? Being able to manipulate machines better? Doing things faster and more efficiently? Or is it keeping the authenticity of the human race while sacrificing efficiency? These are just a few of the questions that can align with the problems that society faces while trying to balance the growing use of AI.