BASIC PROMPT ENGINEERING WITH CHATGPT: AN INTRODUCTION

MODULE 1: Understanding language models and their limitations

OVERVIEW

In this module, we will cover a basic introduction to Large Language Models (LLMs), which are the engines that drive chatbot interfaces like ChatGPT. We will explore their underlying concepts, architectures, and training processes. As we navigate this landscape, we will also discuss some of the limitations and challenges that come with using these powerful tools.

Large language models like the ones behind ChatGPT have revolutionized the way we interact with technology, but it's crucial for us to understand both their potential and their pitfalls. We will examine how these models are trained on vast amounts of data, enabling them to generate human-like text based on the patterns they've learned. By comprehending the inner workings of these models, you'll be better equipped to harness their capabilities effectively and responsibly.

Throughout this and subsequent modules, we will address potential biases, ethical considerations, and other challenges associated with using large-scale language models. As we analyze the limitations of these tools, we will discuss strategies for mitigating biases and other unintended consequences. By the end of this module, you will have a solid understanding of large language models and their limitations, laying the foundation for responsible and effective prompt engineering.

LEARNING OBJECTIVES

By the end of this module you will be able to:

1.1. Describe what ChatGPT is
1.2. Describe what a Large Language Model is
1.3. Describe some of the key benefits of ChatGPT and Large Language Models
1.4. Describe some of the key limitations of Large Language Models and ChatGPT
1.5. Apply basic tips and tricks to getting the most out of ChatGPT

FLOW

START HERE

[Intro video]

Welcome to Basic prompt Engineering with ChatGPT: An Introduction! In this course you will develop foundational skills around using and getting the most out of OpenAI's ChatGPT. These skills will be invaluable in any professional setting where ChatGPT and other AI chatbots are used. They will also be transferrable to other AI chatbots.

Before we dive in, please take a few minutes to watch the video above and read through the following material – you'll find it extremely helpful in preparing for the course and orienting yourself to the rapidly developing area of large language models and AI chatbots.

In the next sections we'll cover a basic introduction to ChatGPT and large language models, the course learning objectives, and some essential tips and tricks to using ChatGPT.

And at any point, if you have questions, or something isn't clear, or you are struggling with an exercise of a concept, please just ask.

Cheers

Andrew Maynard (Instructor)

COURSE OVERVIEW

In November 2022, the company OpenAI publicly released the chatbot "ChatGPT" designed to provide a human interface to the latest version of its powerful large language model (LLM) called GPT (Generative Pre-Trained Transformer). Since then, ChatGPT and the capabilities that it offers has taken the world by storm. These capabilities have grown with ever-more powerful versions of the GPT LLM that is its foundation.

ChatGPT and other LLM-based systems can respond to conversational typed requests in ways that are uncannily human-like. Not only do they draw upon and synthesizing a vast array of human knowledge, but their responses have the structure and nuance of natural human language (or even specialized subsets of it, as in the synthesis of executable program code). The result is systems that can increasingly match humans in tasks ranging from writing letters, essays, and even grant proposals, to summarizing complex ideas in accessible ways, to writing simple code, to creating a personalized and responsive learning environment, and much, much more.

As these systems continue to evolve, there is a growing recognition that platforms like ChatGPT will impact nearly every aspect of our lives over the coming months and years, and transform the ways that people learn, businesses operate, and decisions are made. Although ChatGPT is best known as a stand-alone chatbot, its underlying GPT-4 LLM is now being integrated into other products, from specialized tools for programmers to general-purpose office productivity suites. Consequently, in the future, LLM's may be ubiquitous, and being able to make use of them may be a key factor in ensuring someone's success in the workplace. One way to start preparing for that future is to learn to master the use of tools like ChatGPT today.

This mastery is increasingly being referred to as "prompt engineering." While there are different definitions of what prompt engineering is – including some that refer to the process of developing LLMs rather than using them – in this course prompt engineering is defined as:

> "The art and skill of crafting, optimizing, and employing every-day language to effectively harness the power and capabilities of Large Language Models and AI chatbots, leveraging their potential and making them useful and accessible across a very wide range of professional and personal situations while ensuring accurate and valuable outcomes."

Curiously, the compelling human-like outputs of ChatGPT do not emerge from some deliberate, human-like artificial intelligence under its hood. The LLM within ChatGPT is the natural extension of text-prediction technologies that have been developed over the past few decades. The quality of the outputs of ChatGPT are primarily a product of (a) the very large amount of human text data that its LLM has been trained with to accurately predict responses; and (b) cleverly constructed "prompts" that give enough context to convert a question answering session into a text-completion task. As a result, there is rapidly growing interest in how to get the most out of systems like ChatGPT through the sophisticated use of these prompts. This has led to expertise in "prompt engineering" moving from the domain of coders to people who can formulate, craft, and refine written prompts that maximize the power LLMs.

Between November 2022 and April 2023 online searches for "prompt engineering" increased approximately 20-fold. This was accompanied by companies starting to advertise positions requiring verbal rather than code-based prompt engineering experience.

This course responds directly to this growing expectation that employees in all sectors will have fluency with crafting and engineering prompts for ChatGPT and similar systems. It is built around six modules that introduce the essentials of large language models and associated chatbots; develop skills around basic prompt engineering, including formulation and refinement, generalization and templates, and evaluation and metrics; and explore emerging trends and applications; and contextualize prompt engineering within broader conversations around the broader implications of LLMs.

Through it, you will use ChatGPT to develop and refine your skills around prompt engineering while becoming familiar with how to get the most out of the platform within a variety of contexts. By the end of it, you will have mastered basic skills in prompt engineering.

For more on the approach to prompt engineering used in this course, see "This is not your "traditional" prompt engineering!"

CORE SKILLS/LEARNING OUTCOMES

This course focuses on six core skills/learning outcomes that are important for prompt engineering and using ChatGPT in professional settings. You will refine many of these skills across multiple modules, but each module has a specific focus:

1. **Understanding large language models and their limitations (Module 1):** Through this course, you will become familiar with the underlying concepts of large language models, including how they are trained, their architecture, and their limitations. This will include discussing potential biases, pitfalls, and ethical considerations when using large language models like the one powering ChatGPT.
2. **Prompt formulation and refinement (Module 2):** Through the course, you will learn how to write clear, concise, and unambiguous prompts that effectively communicate the desired information or task. You will also develop the skills of iterating and refining prompts to optimize the model's response, and you will experiment with different phrasings, context, and constraints to achieve the desired outcome.
3. **Developing and using prompt templates (Module 3):** Through the course, you will learn how to create generalized prompt templates that work across a variety of situations, domains, or problems. This skill is essential for maximizing the usefulness of ChatGPT in diverse applications. You will also learn how to identify common structures and adapt their templates accordingly.
4. **Prompt and response evaluation (Module 4):** Through the course, you will become familiar with the process of evaluating language model responses based on various criteria, such as relevance, coherence, and accuracy. You will also develop skills around quantitative and qualitative evaluation methods, and how to develop appropriate metrics for their specific use cases.
5. **Awareness of emerging trends and innovative uses of ChatGPT (Module 5):** Through the course, you will develop an awareness of cutting edge trends in LLM and ChatGPT-like technologies and interfaces, and their applications across multiple areas.
6. **Responsible innovation and broader societal implications of LLMs (Module 6):** Through the course, you will gain an understanding of the wider impact of LLMs like ChatGPT on society, including their potential benefits, risks, and ethical concerns. This will involve topics such as privacy, misinformation, automation, principled innovation, and responsible development and deployment. By cultivating awareness of these broader implications, you will be better equipped to make informed decisions when working with LLMs.

As you work through the course, you will notice that many of the exercises and assignments have been designed to both illustrate different aspects of prompt formulation and engineering, and to help you develop your prompt engineering "muscles."

TIPS AND TRICKS:

Before we start, it's useful to know some very basic tips and tricks for using ChatGPT:

1. Make sure you select "GPT-4" from the model dropdown menu before you start a new session with ChatGPT (only available in ChatGPT Plus)
2. Be aware that when using the GPT-4 model, ChatGPT limits the number of requests you can make over a given time period. Spread your work out so that you don't find yourself out of time when a critical assignment is due!
3. Write your prompts naturally, as if you were speaking to someone else. ChatGPT is trained to understand how you write and "speak" as a person. Don't sweat your prompts, but simply clarify what you are asking or try a different tack if ChatGPT doesn't understand.
4. If you don't know where to begin, just start writing stuff – even notes – and see how ChatGPT responds, then use this to refine your prompts. You can even ask ChatGPT for help with your prompts.
5. Have conversations with ChatGPT.
6. You don't have to be polite when using ChatGPT, and some people suggest you shouldn't as it uses up words and anthropomorphizes the technology. However, I would advocate for being polite and personable as this is a reflection of how you engage with others. Plus, you'll feel much better having a personable conversation with ChatGPT!
7. Sometimes you want to break up your prompts into paragraphs. If you are typing directly into ChatGPT start a new paragraph with the combined keystroke [shift] [return]
8. ChatGPT responses are sometimes terminated before they have finished. If this happens, simply type something like "please continue" and ChatGPT will continue where it left off.
9. If you need to enter a lot of text into a ChatGPT prompt, you will sometimes hit the Chatbot's character limit. In this case, you can break the text down into manageable chunks and let ChatGPT know that you will be uploading it in sections. At the end of each chunk of text, let ChatGPT know there is more to come, and after the last chunk, let ChatGPT know that you have uploaded all of the text. This is a workaround to uploading large amounts of information. However, ChatGPT has a tendency to lose the plot sometimes when provided with a lot of information, so beware!
10. If you get an error message like "there was an error in generating a response" simply try again or (in the worst case) start a new session. This just means there was a glitch in the system.
11. Remember that ChatGPT has no in-built sense of what is true or false, and so it can sometimes provide information that is misleading or just plain wrong. Always use your critical thinking skills to assess responses.

12. Remember that information you provide to ChatGPT may be available to others. It's worth taking the time to read OpenAI's Terms of Use, Data Usage FAQ, and Privacy policy.
13. As was noted in Module 0, the easiest way to document your ChatGPT sessions is to use Chrome and install an extension like ShareGPT or Save ChatGPT.

MODULE OVERVIEW

Intro blurb (above)

Intro video:

- Introduce large language models: result of years of research into machine learning using deep learning and the compute power of Graphical Processor Units, or GPUs
- Have foundations in text prediction, but surprised researchers by sophistication of outputs
- LLMs work by using vast amounts of data to predict the statistically most likely response to an input that a human would give
- The result is an uncanny human-ness to responses, although they are just machines.
- However, the technology is transformative. Since being launched in November 2022, ChatGPT has swept the world, impacting education, business, administration, research, and much more.
- Make sure mention that GPT stands for Generative Pretrained Transformer
- In this module you will work through a series of assignments to develop a better understanding of LLMs and ChatGPT, including their potential and their potential downsides.
- As will all modules, you will be developing your ChatGPT muscles through the exercises, as well as learning directly about ChatGPT and prompt engineering.

COURSE OVERVIEW

Spend an hour simply playing with ChatGPT. Use the Tips and Tricks to get used to the interface

EXERCISE: ChatGPT Orientation and play (20 point)

Spend an hour simply playing with ChatGPT. Use the Tips and Tricks to get used to the interface and how ChatGPT works.

Submit an example of a particularly interesting conversation you had with ChatGPT. This can be about anything.

You should spend no more than an hour on this exercise.

You will be given full points for submitting documentation that demonstrates you have spent an appropriate amount of time and effort on this exercise. Points will be given on submission. Points may be removed at a future date if it appears that you did not spend as much time and effort as expected on the exercise.

EXERCISE: Large Language Model exploration (30 point)

This exercise is designed to help you learn about Large Language Models (LLMs) and LLM chatbots like ChatGPT through your own research.

In the first part you will use Google (or similar) to research papers, articles, blogs, podcasts etc. on LLMs and ChatGPT and use these to develop a better understanding of what they are and their limitations.

In the second part you will further refine your understanding of LLMs through a conversation with ChatGPT

The third part is a simple ChatGPT-conducted quiz that tests your understanding of LLMs.

The whole exercise should take no more than 1 – 2 hours.

Background Research (10)

1. Spend some time exploring online sources to learn about the development, uses, and limitations of LLMs and ChatGPT.
2. Provide links to three sources you found useful, and say in 2 – 3 sentences why you found them useful. Do not use ChatGPT to construct these sentences.
3. Do **not** use ChatGPT to construct these sentences

You will be given full points for completing this exercise. Points may be removed at a future date if it appears that you did not spend as much time and effort as expected on the exercise.

ChatGPT Conversation (10)

1. Have a conversation with ChatGPT about what LLMs are, what GPT stands for, and the benefits and limitations of LLMs and GPTs. Your conversation should consist of at least 10 prompts, but can be longer than this.
2. Submit documentation of your conversation to the assignment (see the guidelines on how to do this in Module 0)

You will be given full points for completing this exercise. Points may be removed at a future date if it appears that you did not spend as much time and effort as expected on the exercise.

LLM and GPT quiz (10)

1. Start a new chat with ChatGPT (making sure you are using GPT-4) and cut and paste the following prompt: “Hi ChatGPT. My name is [add your full name] and I am in a class where we are learning about the uses and limitations of LLMs and GPTs. Please ask me three simple questions about the uses and limitations of LLMs to test my understanding. After each question, please wait for my answer before asking the next one. When you have all three of my answers, please provide an assessment of how good they are, and give me a grade from A to C.”
2. You can repeat this quiz as many times as you want. You should also treat it as an “open book” (or “open Google”) quiz.
3. When you are happy with your responses, submit documentation of the ChatGPT below. By default you will get full marks for this assignment once your documentation is submitted (we will not take account of the grade that ChatGPT gives you on this occasion – this is for your own use). However we will be checking submissions, and will adjust points if you appear to be struggling to understand the questions or have not put sufficient effort into this exercise.

EXERCISE: Error, bias, and other failure modes in LLMs (30 point)

This exercise is designed to help you better understand some of the errors and biases that can be present in conversations with ChatGPT. It only scratches the surface of errors and bias within LLMs but is useful in understanding how to spot and navigate them.

Bias in AI, including in LLMs, is a very real concern. Chatbots like ChatGPT reflect the biases in their training data, which in turn reflects human biases that span gender and ethnicity biases to political and socioeconomic biases and more.

However, as LLMs like ChatGPT evolve, efforts are being made to address and minimize these biases, meaning that detecting and analyzing bias is a moving target.

Errors are also a significant concern. LLMs are known to “hallucinate” where they present factually incorrect (and often made-up) information with authority.

This exercise uses ChatGPT to explore hallucinations, bias, and strategies for navigating both.

The whole exercise should take no more than 1 – 2 hours.

Hallucination demonstration (10)

1. Open a new ChatGPT session using GPT4. Enter the following prompt: "Hi ChatGPT. Please give me citations for three papers that explain the concept of hallucinations with LLMs"
2. Check the validity of the citations by following any links that are provided, and doing Google searches
3. Have a conversation with ChatGPT about any errors you find.
4. Submit documentation of your conversation below.

You will be given full points for completing this exercise. Points may be removed at a future date if it appears that you did not spend as much time and effort as expected on the exercise.

Bias (10)

1. Have a conversation with ChatGPT about bias. Push ChatGPT to expand on different types of bias, and explore strategies for identifying bias in responses and creating prompts that reduce the chances of bias.
2. Submit documentation of your conversation below.

You will be given full points for completing this exercise. Points may be removed at a future date if it appears that you did not spend as much time and effort as expected on the exercise.

Failure Modes (10)

1. Start a new ChatGPT session with GPT-4 and ask "Hi ChatGPT. What are the most significant failure modes of ChatGPT that I should be aware of?"
2. Continue the conversation to further develop your understanding of failure modes. Where appropriate, use Google to check what ChatGPT is telling you.
3. Submit documentation of your conversation below.

You will be given full points for completing this exercise. Points may be removed at a future date if it appears that you did not spend as much time and effort as expected on the exercise.

ASSIGNMENT: Reflection (20 point)

1. Post a short reflection (100 – 200 words) on what you have learned about the limitations of ChatGPT and how to ensure these don't unduly impact your work. Do not use ChatGPT to write your reflection.

Reflections will get between 18 – 20 points unless it is clear that little effort has been made or ChatGPT has been used.