

ChatGPT Workshop for Data Engineers

8 academic hours

Overview:

The workshop provides a comprehensive overview of the OpenAI API and its utilization. It explores the roles and definitions within the system, user, assistant, and function context. The workshop delves into the capabilities of the OpenAI API, including natural language processing, language translation, chatbot development, and content generation.

One of the key focuses of the workshop is to explain the significance of the OpenAI API in empowering developers to leverage powerful language models. Participants will learn about API access and usage, covering the registration and authentication process, obtaining API keys, and making API calls. The workshop also provides detailed explanations of different API endpoints and their functionalities, such as generating text, translating languages, and analyzing sentiment. Practical examples and code snippets are included to demonstrate the practical usage of the API.

Participants will gain an in-depth understanding of working with API parameters, exploring various options available when making API requests. Parameters like temperature, max tokens, and top p/top k values will be discussed, along with their effects on the generated output. The workshop will guide participants on how to adjust these parameters to influence creativity, length, and coherence of the generated text. Best practices and guidelines for selecting appropriate parameter values will also be shared.

The workshop will address the limitations and considerations when working with the OpenAI API. Topics such as rate limits, cost considerations, and responsible and ethical usage will be highlighted. Participants will learn about potential challenges in managing bias in generated content and mitigating risks associated with malicious use. OpenAI's guidelines and recommendations for safe and ethical utilization of the API will also be discussed.

The workshop includes an overview of the Python API, focusing on managing conversations using code and separating sessions. Participants will gain insights into the integration layer needs and how to manage authentication and authorization. Topics like ad-hoc session, OpenAI private environment, and the usage of Vector DB will be explored to provide participants with a comprehensive understanding of the OpenAI API ecosystem.

By the end of the workshop, participants will have the knowledge and skills to effectively access, utilize, and customize the OpenAI API for various applications, enabling them to leverage the power of language models in their projects.

Who Should Attend:

This workshop is relevant for data engineers and big data developers from all disciplines, who are seeking to better understand ChatGPT and its capabilities for assisting with everyday tasks.

Required Skills:

- Experienced Data Engineers and BI Developers or any other Data related field with a solid background in SQL or Python

Course Contents:

Section 1: Intro to ChatCPT and Generative AI

- Generative AI
- Intro to Language Models
- Intro to ChatGPT
- Chat Terminology

Section 2: Understanding OpenAI API

- Introduction to OpenAI API
 - Overview of OpenAI API
 - Explaining what it is and how it can be utilized
 - Role's definition: system, user, assistant, function
 - Discuss the capabilities of the API, such as natural language processing, language translation, chatbot development, and content generation
 - Explain the significance of OpenAI API in enabling developers to leverage powerful language models
- API Access and Usage
 - Cover the process of accessing and using the OpenAI API
 - Discuss the registration and authentication process, obtaining API keys, and making API calls
 - Explain the different API endpoints and their functionalities, such as generating text, translating languages, or analyzing sentiment
 - Provide examples and code snippets to demonstrate the practical usage of the API
- Working with API Parameters
 - Dive into the various parameters and options available when making API request
 - Discuss parameters such as temperature, max tokens, and top p/top k values, explaining their effects on the generated output
 - Explore how adjusting these parameters can influence the creativity, length, and coherence of the generated text
 - Discuss best practices and guidelines for selecting appropriate parameter values
- API Limitations and Considerations
 - Highlight the limitations and considerations when working with the OpenAI API
 - Discuss factors such as rate limits, cost considerations, and the importance of responsible and ethical usage
 - Address potential challenges such as managing bias in generated content and mitigating risks associated with malicious use
 - Discuss OpenAI's guidelines and recommendations for safe and ethical utilization of the API

Section 3: Python API and integration layer definition

- Python API
 - Overview of Python API
 - Manage a conversation using code
 - Separation of sessions – managed by code
- Integration layer

- Overview of integration layer needs
- Managing authentication and authorization
- Ad-hoc session
- OpenAI private environment
- Why and how to use Vector DB