

Exploring the OpenAI API



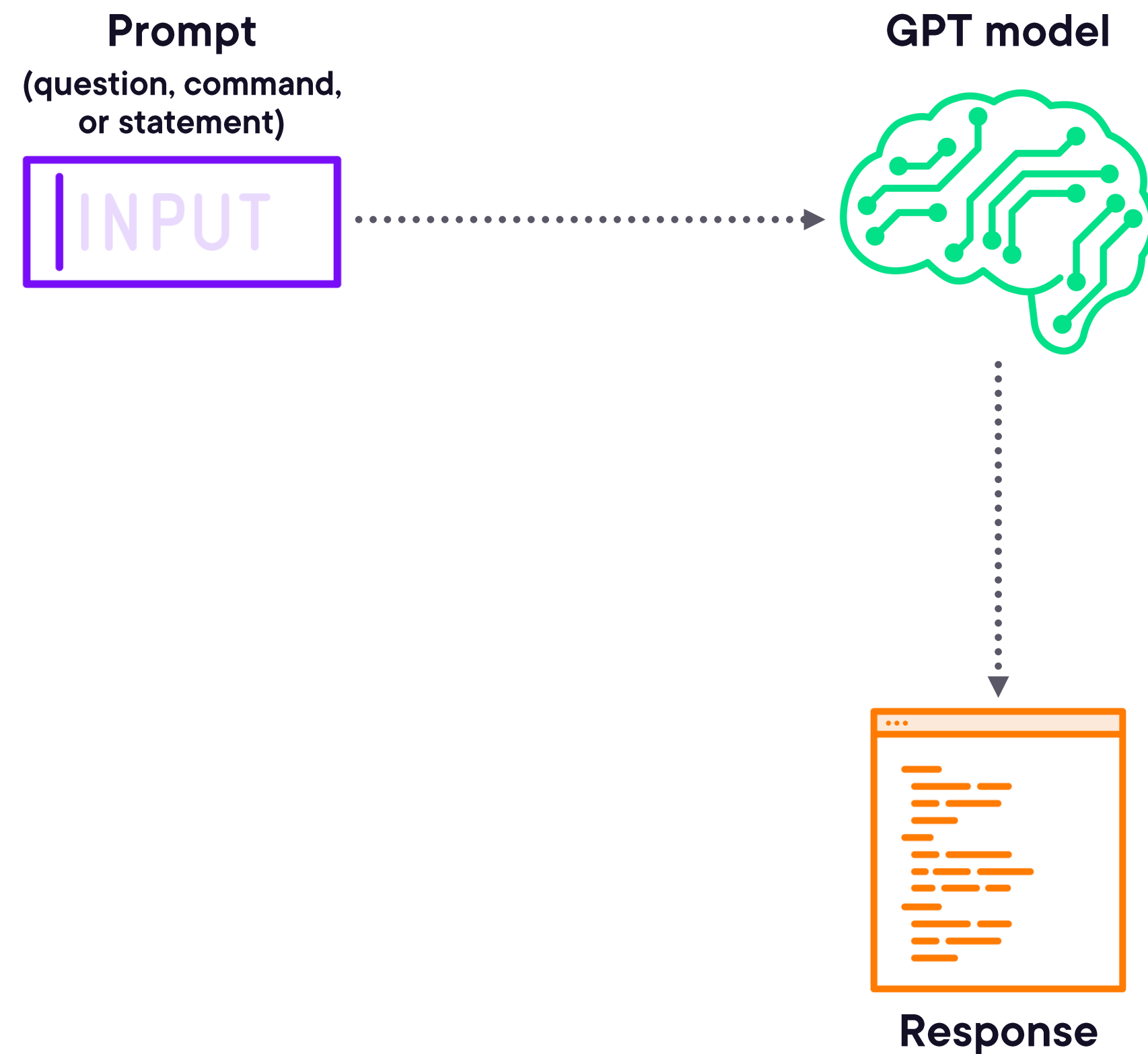
Esteban Herrera

Author

@eh3rrera | eherrera.net



Interacting with GPT Models



**How we write the prompt
plays an important part in
the quality of the response
we get from the model.**



Predicting Next Words

You are a helpful and friendly programmer that...
...strives to use my programming knowledge
and understanding to assist you with any
problems you might be facing.
So, how may I assist you today?



Predicting Next Words Tokens

You are a helpful and friendly programmer that...

...strives to use my programming knowledge

and understanding to assist you with any

problems you might be facing.

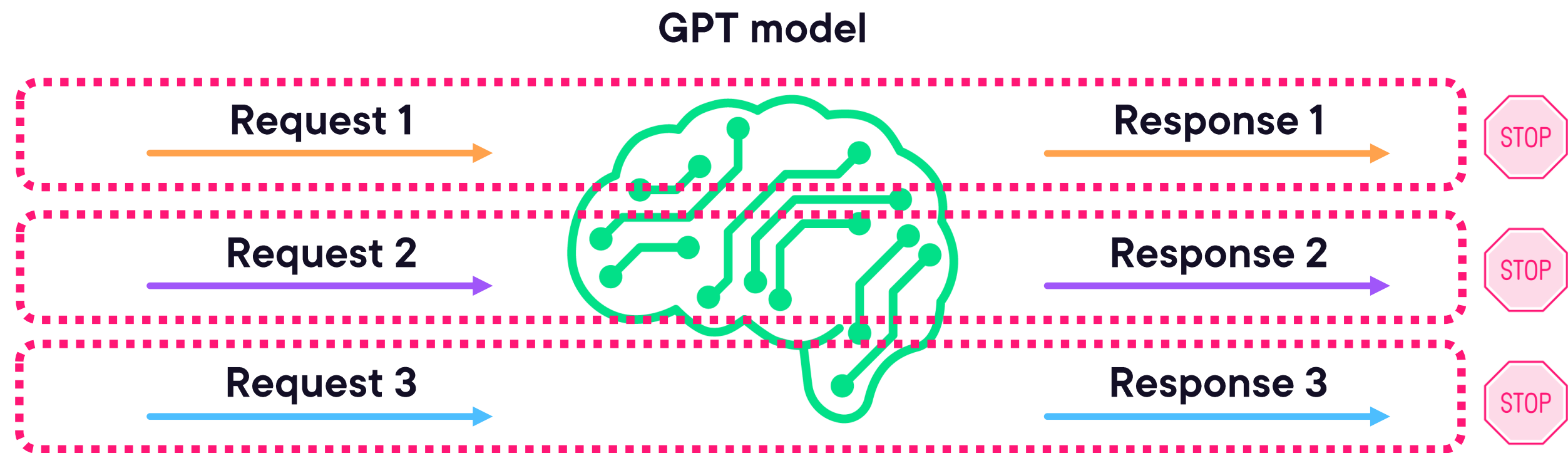
So, how may I assist you today?



Understanding Token IDs



Stateless Interactions



Context Window

**GPT-3.5
4K**

GPT, or Generative Pretrained Transformer, models are a series of language prediction models characterized by their size and their capacity to generate human-like text based on a given input.

GPT models are a part of a broader family of machine learning models known as transformers. Transformers were introduced in a 2017 paper called "Attention is All You Need" by Vaswani et al. The key innovation of transformers is the attention mechanism, which allows the model to weigh the importance of different words in a sentence when generating predictions.

GPT models work by predicting the next word in a sequence. This is called autoregressive language modeling. For example, if you feed the model the sentence "The cat sat on the...", it would predict the next word as "mat" or "couch" or some other plausible word.



Context Window

GPT, or Generative Pretrained Transformer, models are a series of language prediction models characterized by their size and their capacity to generate human-like text based on a given input.

←..... Out of the context window

**GPT-3.5
4K**

GPT models are a part of a broader family of machine learning models known as transformers. Transformers were introduced in a 2017 paper called "Attention is All You Need" by Vaswani et al. The key innovation of transformers is the attention mechanism, which allows the model to weigh the importance of different words in a sentence when generating predictions.

GPT models work by predicting the next word in a sequence. This is called autoregressive language modeling. For example, if you feed the model the sentence "The cat sat on the...", it would predict the next word as "mat" or "couch" or some other plausible word.





OpenAI's API Models



Snapshot Versions

gpt-4-0613

gpt-3.5-turbo-0613

The diagram illustrates the structure of snapshot version numbers. For 'gpt-4-0613', the '06' represents the month and '13' represents the day. For 'gpt-3.5-turbo-0613', the '06' represents the month and '13' represents the day. Dotted lines with arrows point from the corresponding digits to the labels 'Month' and 'Day'.

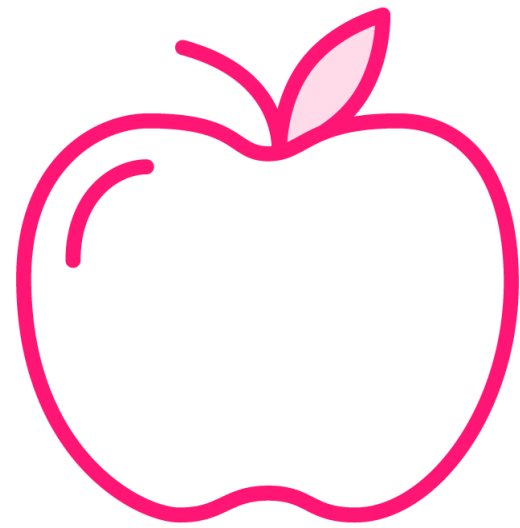
Deprecated about 3 months after a new version is released



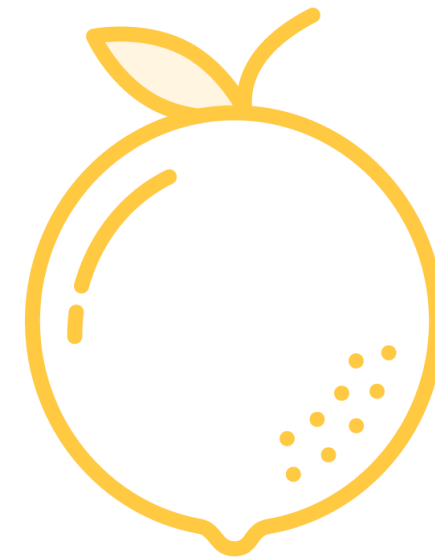
Snapshot Versions



Embedding



(sweetness, color)
(7, 8)



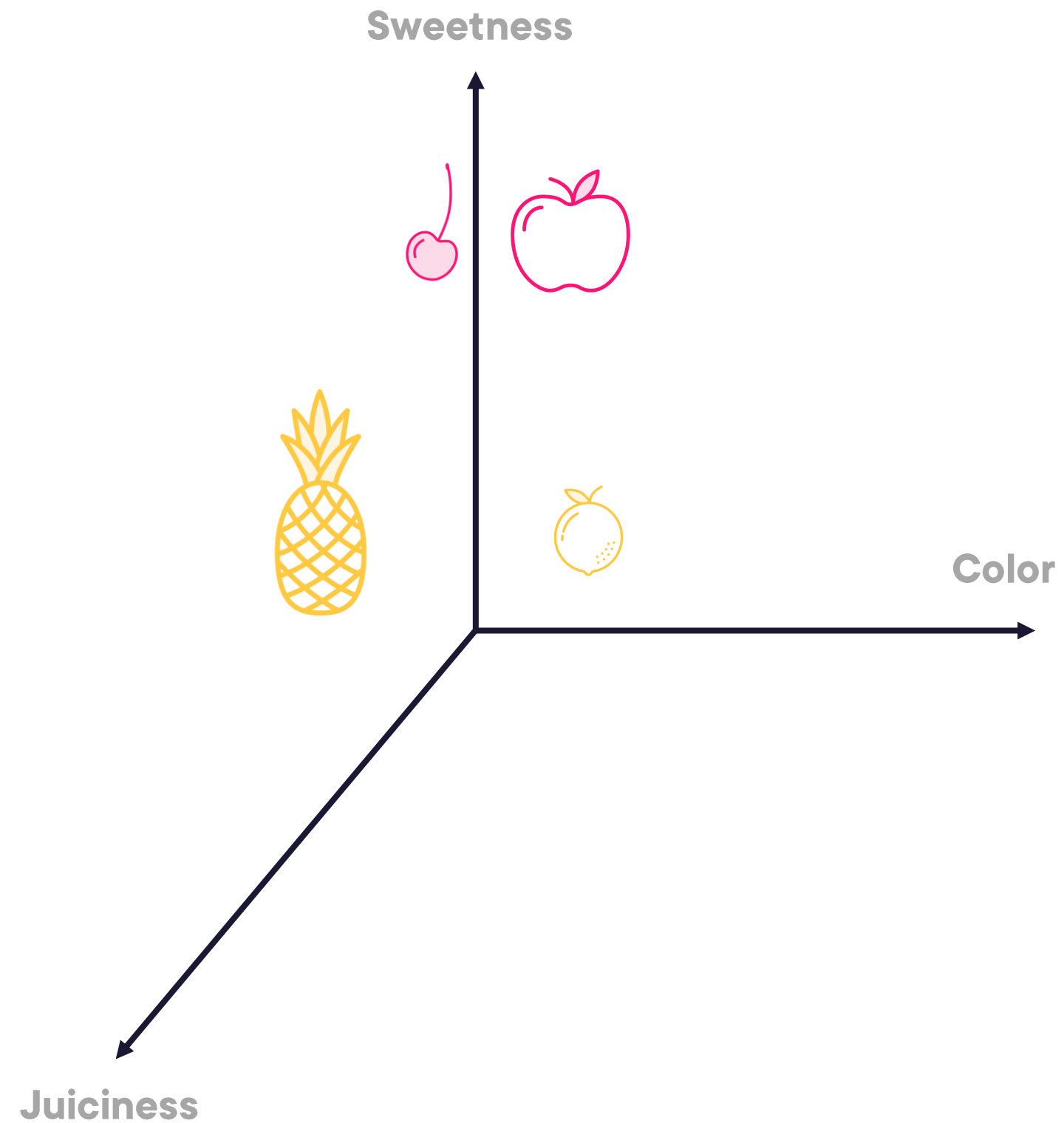
(sweetness, color)
(2, 5)

Embedding Dimensions

Semantic relationships

Similarity

Classification



Setting up the Python OpenAI Library



Virtual Environment

A directory that contains a complete Python installation.




```
python3 -m venv myenv
```

Create a Virtual Environment

Use the built-in venv module



```
# Activate environment
myenv\Scripts\Activate # On Windows
source myenv/bin/activate # On Mac or Linux

# Deactivate environment (regardless of the system)
deactivate
```

Use a Virtual Environment





The Images API



On Windows

```
setx OPENAI_API_KEY "YOUR_API_KEY"
```

On Mac or Linux

```
export OPENAI_API_KEY=YOUR_API_KEY
```

Setting an Environment Variable

Not recommended, the key could be recorded in your command-line history



Better Option for Setting an Environment Variable on Windows

1. Press Windows key + R to open the Run dialog
2. Type sysdm.cpl and press Enter. This will open the System Properties window
3. Go to the Advanced tab and click on Environment Variables
4. Under the User variables section, click on New
5. Enter OPENAI_API_KEY as the variable name and the respective API key as the variable value
6. Click OK to save it



Better Option for Setting an Environment Variable on Mac or Linux

Locate the shell profile file

- ~/.bash_profile (for Bash)
- ~/.zshrc (for ZSH)

Add the export command there

- export OPENAI_API_KEY=YOUR_API_KEY

Optionally, reload your shell profile file to apply the change immediately

- source ~/.bash_profile
- source ~/.zshrc





The Audio API



Important Things to Consider



The size limit of the audio file



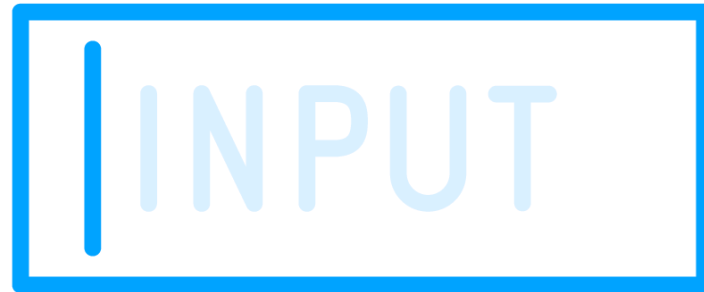
The supported input file types



The supported languages



Using Prompts



Setting the technical context

- "In this course, we will cover concepts such as neural networks, backpropagation, and convolutional neural networks (CNN)."

Ensuring capitalization and punctuation

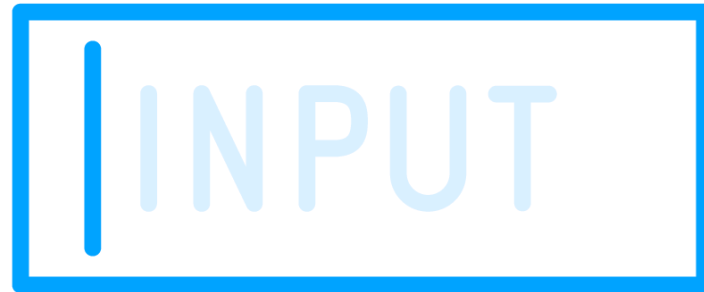
- "Welcome to the course! We'll be exploring the fascinating world of AI. Let's dive in."

Preserving context in segmented files

- "In the previous lesson, we discussed some AI concepts. Now, let's talk about GTP models."



Using Prompts



Including filler words for informal tone

- "So, um, today we're, like, going to talk about something super cool, uh, the human brain!"

Handling foreign language styles

- Traditional Chinese: "歡迎參加本課程。今天，我們將學習中國的歷史。"
- Simplified Chinese: "欢迎参加本课程。今天，我们将学习中国的历史。"

Clarifying acronyms and abbreviations

- "Today we will talk about DNA and RNA, and their roles in biology and agriculture."





The Chat Completions API



Chat Messages

```
[  
  {"role": "system", "content": "You are a Python programmer."},  
  {"role": "user", "content": "Hi!"},  
  {"role": "assistant", "content": "Hi! How can I help you?"}  
]
```



Chat Messages

```
[  
  {"role": "system", "content": "You are a helpful assistant."},  
  {"role": "user", "content": "Hi!"},  
  {"role": "assistant", "content": "Hi! How can I help you?"}  
]
```



Chat Messages

```
[  
  {"role": "system", "content": "You are a helpful assistant."},  
  {"role": "user", "content": "Example of a positive course review: "},  
  {"role": "assistant", "content": "The 'Mastering Python' course was simply outstanding"},  
  {"role": "user", "content": "Example of a negative course review: "},  
  {"role": "assistant", "content": "The 'Learning Python' course was was quite underwhelming"},  
  {"role": "user", "content": "Give me a negative review for the 'Python Basics' course: "},  
  {"role": "assistant", "content": "The 'Python Basics' course was was quite underwhelming"}  
]
```





API Pricing



Summary



Key terms

- Prompts
- Tokens
- Context window

API endpoints

- Language
- Embeddings
- Image generation
- Text-to-speech
- Speech-to-text
- Moderation



Summary



OpenAI official Python library

- Image generation
- Speech recognition (with a prompt)
- Answer C# related questions (chat completions)

Pricing

- Number of used tokens
- Images generated
- Minutes of transcribed audio
- Characters converted into audio



Up Next:

Using the Chat Completions API

