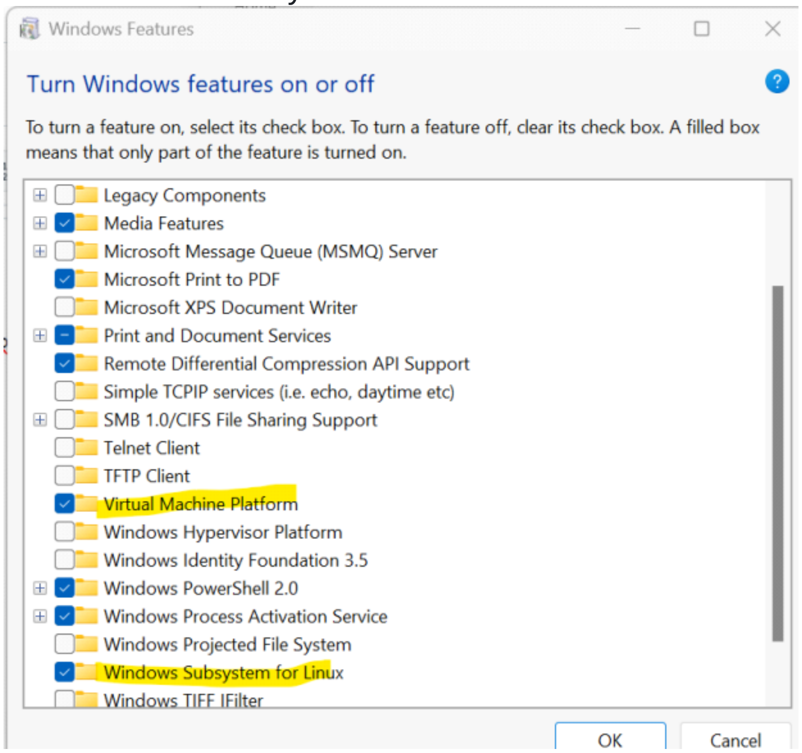
Start WSL features

Thursday, December 14, 2023 11:24 PM

Turn on window feature:

- Virtual Machine Platform
- Windows Subsystem for linux



Turn Windows features on or off

To turn a feature on, select its check box. To turn a feature off, clear its check box. A filled box means that only part of the feature is turned on.

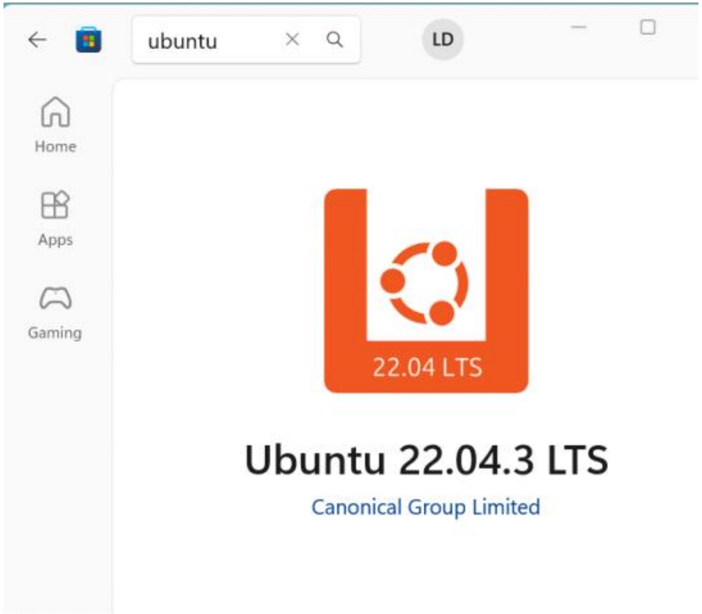
- ⊞ □ Legacy Components
- ⊞ ☑ Media Features
- ⊞ □ Microsoft Message Queue (MSMQ) Server
- ☑ Microsoft Print to PDF
- □ Microsoft XPS Document Writer
- ⊞ ▣ Print and Document Services
- ☑ Remote Differential Compression API Support
- □ Simple TCPIP services (i.e. echo, daytime etc)
- ⊞ □ SMB 1.0/CIFS File Sharing Support
- □ Telnet Client
- □ TFTP Client
- ☑ Virtual Machine Platform
- □ Windows Hypervisor Platform
- □ Windows Identity Foundation 3.5
- ⊞ ☑ Windows PowerShell 2.0
- ⊞ ☑ Windows Process Activation Service
- □ Windows Projected File System
- ☑ Windows Subsystem for Linux
- □ Windows TIFF IFilter

From <https://github.com/phaledang/ai-python/wiki/Start-WSL2>

Install Ubuntu from windows store

Thursday, December 14, 2023 11:25 PM

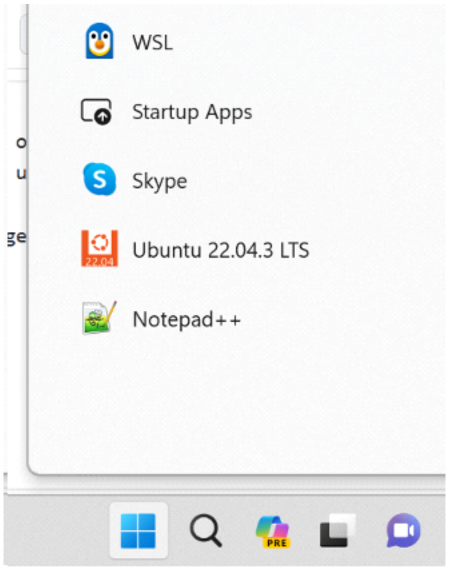
On windows store, search and install the Ubuntu



From <https://github.com/phaledang/ai-python/wiki/Add-Ubuntu>

Run wsl

Thursday, December 14, 2023



From start, click WSL or the Ubuntu name to open the ubuntu

Add Docker on Ubuntu

Thursday, December 14, 2023 11:26 PM

Run below script on Ubuntu

```
sudo apt-get update

sudo apt-get install
apt-transport-https
ca-certificates
curl
gnupg
lsb-release

curl -fsSL https://download.docker.com/linux/ubuntu/gpg | sudo gpg --dearmor -o
/usr/share/keyrings/docker-archive-keyring.gpg

echo
"deb [arch=amd64 signed-by=/usr/share/keyrings/docker-archive-keyring.gpg]
https://download.docker.com/linux/ubuntu
$(lsb_release -cs) stable" | sudo tee /etc/apt/sources.list.d/docker.list > /dev/null

sudo apt-get update

sudo apt-get install docker-ce docker-ce-cli containerd.io

sudo service docker start

sudo groupadd docker

sudo usermod -aG docker $USER

newgrp docker `

## Check docker version:

docker -v
```

References

- https://jflower.co.uk/install-docker-engine-without-docker-desktop-on-windows/

Start docker on the ubuntu when computer is started

Thursday, December 14, 2023 11:44 PM

Run below command on the ubuntu

```
sudo service docker start
```

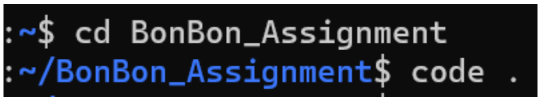
From <https://github.com/phaledang/ai-python/wiki/Start-Docker-on-start-up>

Install miniconda

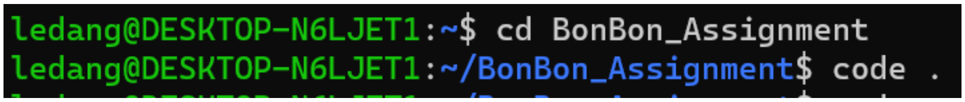
Thursday, December 14, 2023 11:23 PM

Open code

Thursday, December 14, 2023 10:59 PM



```
:~$ cd BonBon_Assignment
:~/BonBon_Assignment$ code .
```



```
ledang@DESKTOP-N6LJET1:~$ cd BonBon_Assignment
ledang@DESKTOP-N6LJET1:~/BonBon_Assignment$ code .
```

Create conda Enviroment

Friday, December 15, 2023   9:12 PM

```
1. `conda create -n bonbon_ass python=3.11`
2. `conda activate bonbon_ass`
3. `pip install -r requirements.txt`

# To activate this environment, use
#
#     $ conda activate bonbon_ass
#
# To deactivate an active environment, use
#
#     $ conda deactivate
```

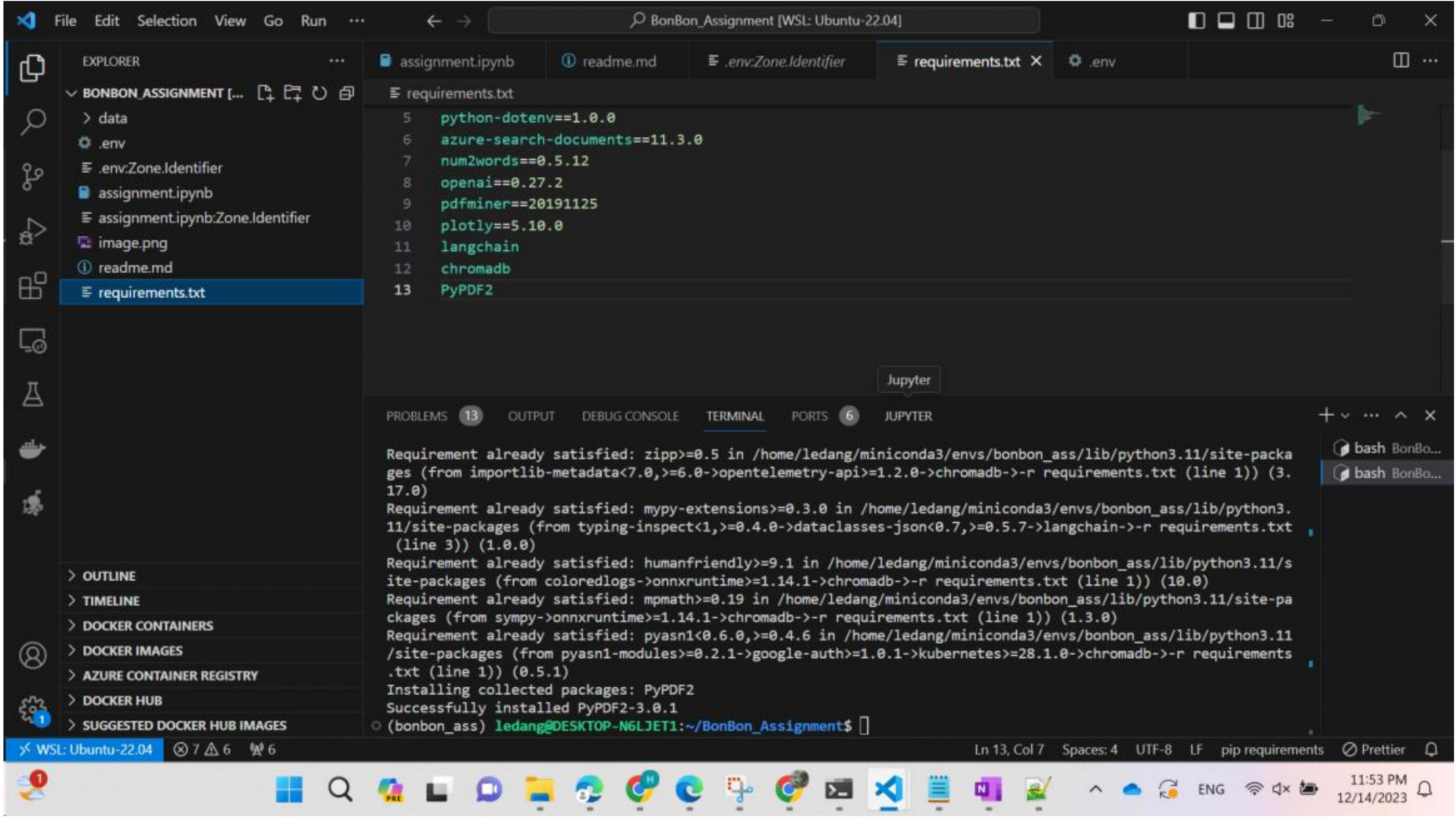
Requirement.txt

Friday, December 15, 2023 9:15 PM

```
chromadb
openai
langchain
langchainhub
python-dotenv==1.0.0
azure-search-documents==11.3.0
num2words==0.5.12
openai==0.27.2
pdfminer==20191125
plotly==5.10.0
langchain
chromadb
PyPDF2
tiktoken
pandas
```

Thursday, December 14, 2023 11:54 PM

Install required package

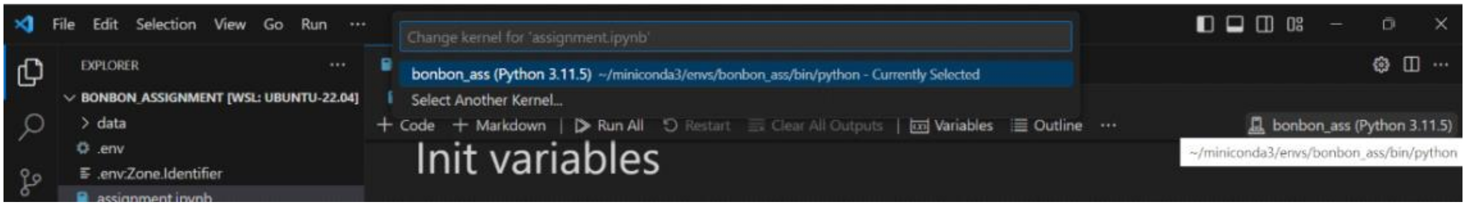


```
python-dotenv==1.0.0
azure-search-documents==11.3.0
num2words==0.5.12
openai==0.27.2
pdfminer==20191125
plotly==5.10.0
langchain
chromadb
PyPDF2
```

```
Requirement already satisfied: zipp>=0.5 in /home/ledang/miniconda3/envs/bonbon_ass/lib/python3.11/site-packa
ges (from importlib-metadata<7.0,>=6.0->opentelemetry-api>=1.2.0->chromadb->-r requirements.txt (line 1)) (3.
17.0)
Requirement already satisfied: mypy-extensions>=0.3.0 in /home/ledang/miniconda3/envs/bonbon_ass/lib/python3.
11/site-packages (from typing-inspect<1,>=0.4.0->dataclasses-json<0.7,>=0.5.7->langchain->-r requirements.txt
 (line 3)) (1.0.0)
Requirement already satisfied: humanfriendly>=9.1 in /home/ledang/miniconda3/envs/bonbon_ass/lib/python3.11/s
ite-packages (from coloredlogs->onnxruntime>=1.14.1->chromadb->-r requirements.txt (line 1)) (10.0)
Requirement already satisfied: mpmath>=0.19 in /home/ledang/miniconda3/envs/bonbon_ass/lib/python3.11/site-pa
ckages (from sympy->onnxruntime>=1.14.1->chromadb->-r requirements.txt (line 1)) (1.3.0)
Requirement already satisfied: pyasn1<0.6.0,>=0.4.6 in /home/ledang/miniconda3/envs/bonbon_ass/lib/python3.11
/site-packages (from pyasn1-modules>=0.2.1->google-auth>=1.0.1->kubernetes>=28.1.0->chromadb->-r requirements
.txt (line 1)) (0.5.1)
Installing collected packages: PyPDF2
Successfully installed PyPDF2-3.0.1
(bonbon_ass) ledang@DESKTOP-N6LJET1:~/BonBon_Assignment$ 
```

Select the kernel

Thursday, December 14, 2023 10:48 PM



Change kernel for 'assignment.ipynb'

bonbon_ass (Python 3.11.5) ~/miniconda3/envs/bonbon_ass/bin/python - Currently Selected

Select Another Kernel...

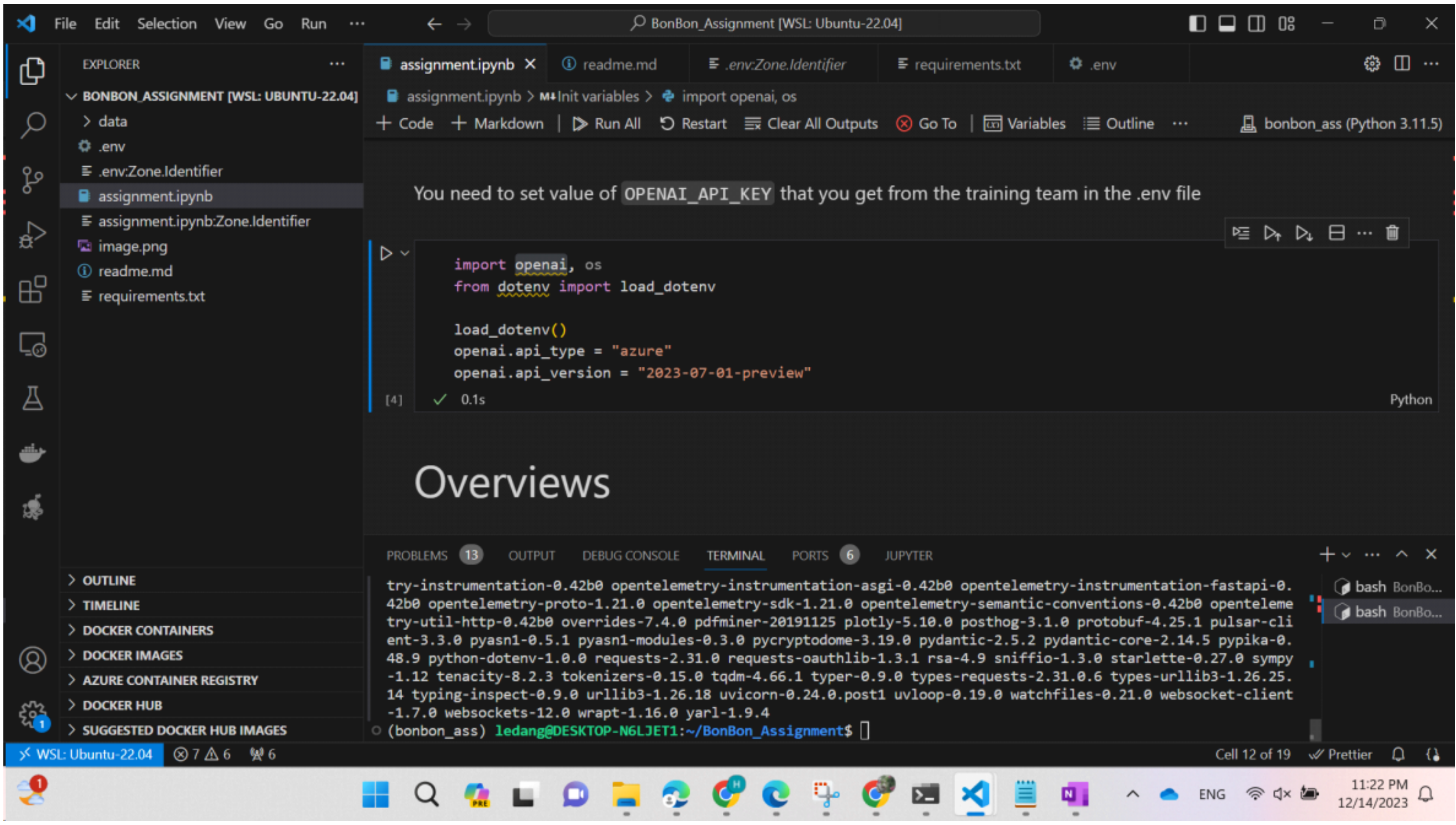
bonbon_ass (Python 3.11.5)

~/miniconda3/envs/bonbon_ass/bin/python

Init variables

Load env

Thursday, December 14, 2023 11:22 PM



File Edit Selection View Go Run ···
BonBon_Assignment [WSL: Ubuntu-22.04]
EXPLORER
BONBON_ASSIGNMENT [WSL: UBUNTU-22.04]
data
.env
.env:Zone.Identifier
assignment.ipynb
assignment.ipynb:Zone.Identifier
image.png
readme.md
requirements.txt
OUTLINE
TIMELINE
DOCKER CONTAINERS
DOCKER IMAGES
AZURE CONTAINER REGISTRY
DOCKER HUB
SUGGESTED DOCKER HUB IMAGES
assignment.ipynb
readme.md
.env:Zone.Identifier
requirements.txt
.env
assignment.ipynb > M↓Init variables > import openai, os
+ Code + Markdown | Run All Restart Clear All Outputs Go To | Variables Outline ···
bonbon_ass (Python 3.11.5)
PROBLEMS 13 OUTPUT DEBUG CONSOLE TERMINAL PORTS 6 JUPYTER
bash BonBo...
bash BonBo...
WSL: Ubuntu-22.04 ⊗ 7 ⚠ 6 6
Cell 12 of 19 Prettier
ENG
11:22 PM
12/14/2023

You need to set value of `OPENAI_API_KEY` that you get from the training team in the .env file

```python
import openai, os
from dotenv import load_dotenv

load_dotenv()
openai.api_type = "azure"
openai.api_version = "2023-07-01-preview"
```

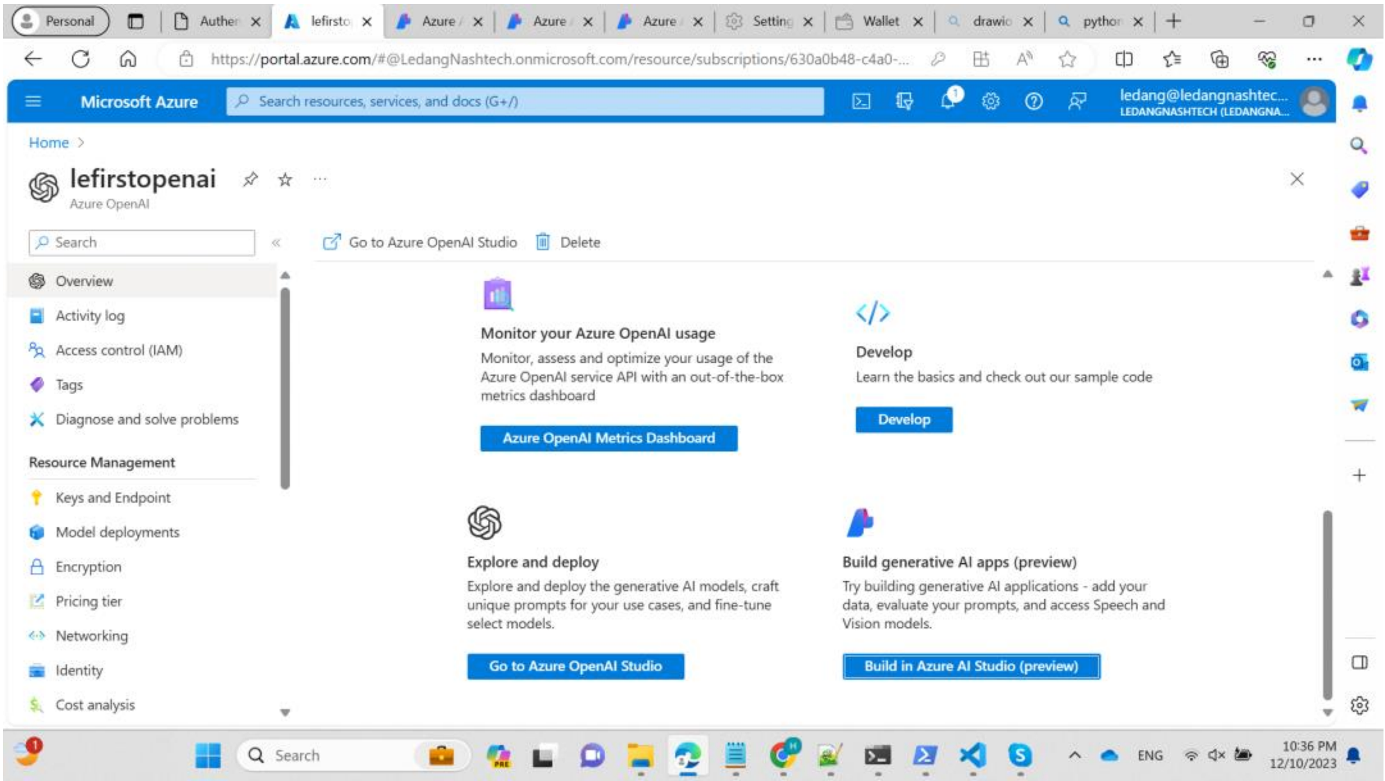
[4] ✓ 0.1s Python

Overviews

```
try-instrumentation-0.42b0 opentelemetry-instrumentation-asgi-0.42b0 opentelemetry-instrumentation-fastapi-0.
42b0 opentelemetry-proto-1.21.0 opentelemetry-sdk-1.21.0 opentelemetry-semantic-conventions-0.42b0 openteleme
try-util-http-0.42b0 overrides-7.4.0 pdfminer-20191125 plotly-5.10.0 posthog-3.1.0 protobuf-4.25.1 pulsar-cli
ent-3.3.0 pyasn1-0.5.1 pyasn1-modules-0.3.0 pycryptodome-3.19.0 pydantic-2.5.2 pydantic-core-2.14.5 pypika-0.
48.9 python-dotenv-1.0.0 requests-2.31.0 requests-oauthlib-1.3.1 rsa-4.9 sniffio-1.3.0 starlette-0.27.0 sympy
-1.12 tenacity-8.2.3 tokenizers-0.15.0 tqdm-4.66.1 typer-0.9.0 types-requests-2.31.0.6 types-urllib3-1.26.25.
14 typing-inspect-0.9.0 urllib3-1.26.18 uvicorn-0.24.0.post1 uvloop-0.19.0 watchfiles-0.21.0 websocket-client
-1.7.0 websockets-12.0 wrapt-1.16.0 yarl-1.9.4
(bonbon_ass) ledang@DESKTOP-N6LJET1:~/BonBon_Assignment$ 
```

Copy config from Azure OpenAI resource to .env file

Saturday, December 2, 2023 3:38 PM


Personal
https://portal.azure.com/#@LedangNashtech.onmicrosoft.com/resource/subscriptions/630a0b48-c4a0-...
Microsoft Azure
Search resources, services, and docs (G+/)
ledang@ledangnashtec...
LEDANGNASHTECH (LEDANGNA...
Home >
lefirstopenai
Azure OpenAI
Search
Go to Azure OpenAI Studio
Delete
Overview
Activity log
Access control (IAM)
Tags
Diagnose and solve problems
Resource Management
Keys and Endpoint
Model deployments
Encryption
Pricing tier
Networking
Identity
Cost analysis
Azure OpenAI Metrics Dashboard
Develop
Go to Azure OpenAI Studio
Build in Azure AI Studio (preview)
Search
ENG
10:36 PM
12/10/2023

Monitor your Azure OpenAI usage

Monitor, assess and optimize your usage of the Azure OpenAI service API with an out-of-the-box metrics dashboard

Develop

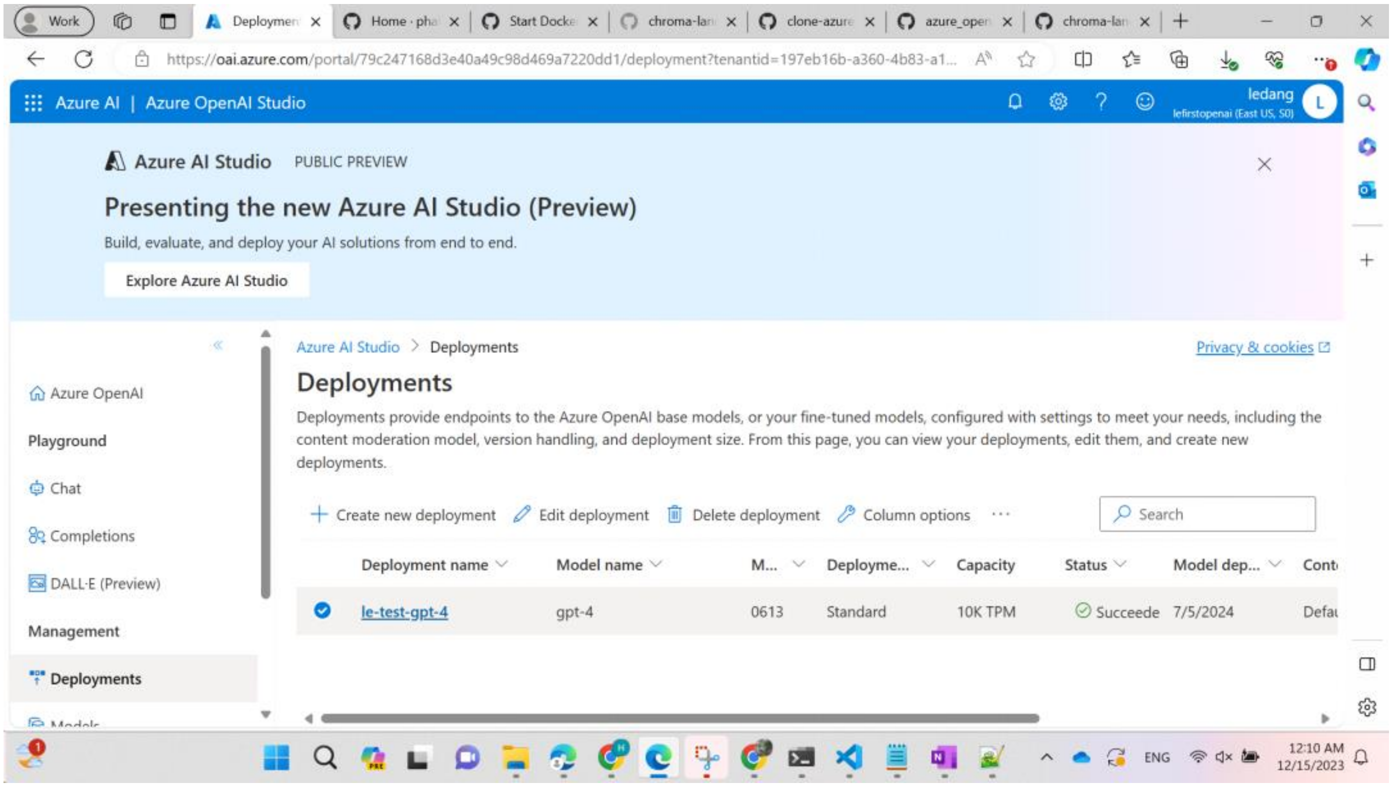
Learn the basics and check out our sample code

Explore and deploy

Explore and deploy the generative AI models, craft unique prompts for your use cases, and fine-tune select models.

Build generative AI apps (preview)

Try building generative AI applications - add your data, evaluate your prompts, and access Speech and Vision models.



Presenting the new Azure AI Studio (Preview)

Build, evaluate, and deploy your AI solutions from end to end.

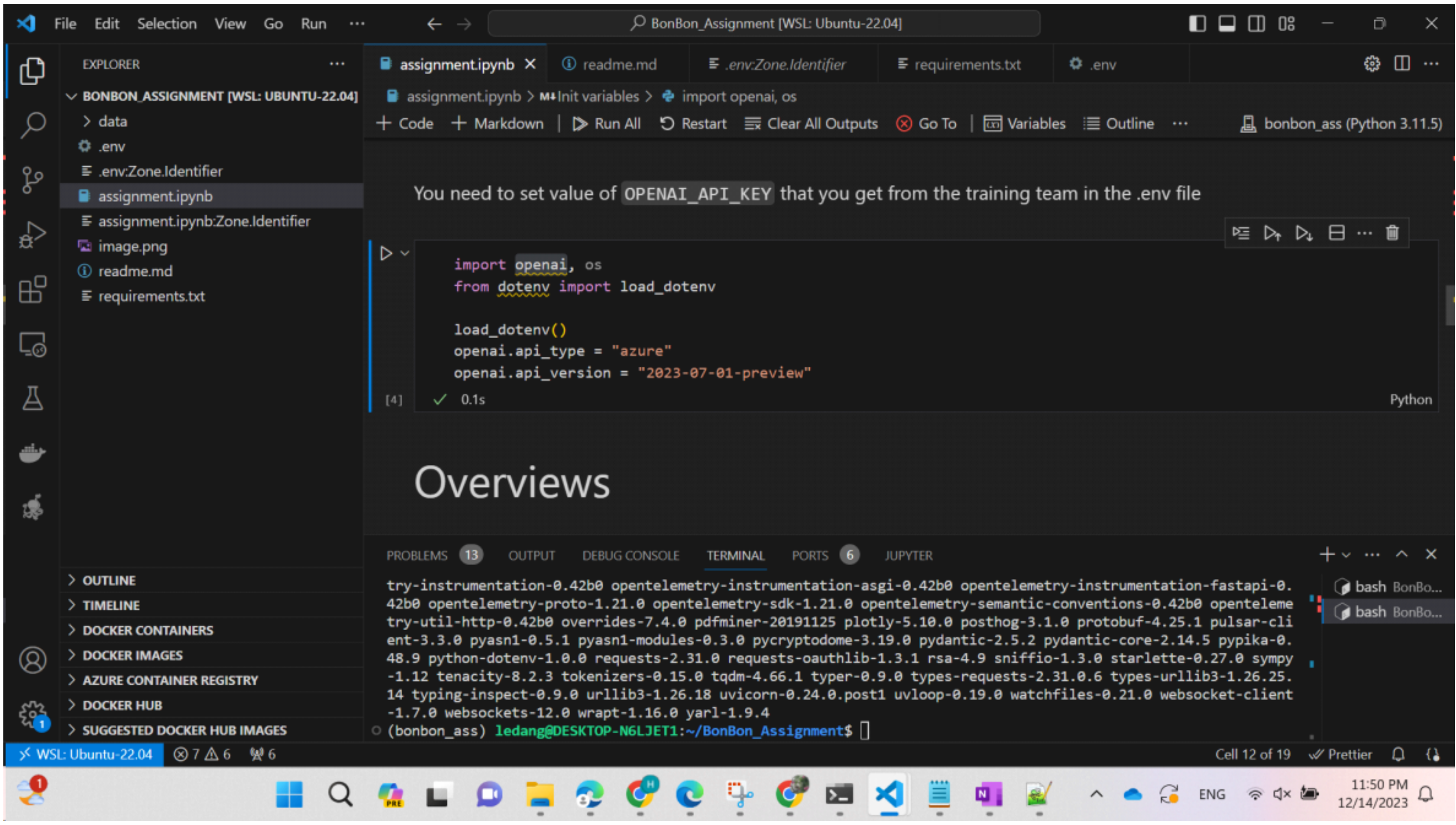
Deployments

Deployments provide endpoints to the Azure OpenAI base models, or your fine-tuned models, configured with settings to meet your needs, including the content moderation model, version handling, and deployment size. From this page, you can view your deployments, edit them, and create new deployments.

Deployment name	Model name	M...	Deployme...	Capacity	Status	Model dep...	Cont
le-test-gpt-4	gpt-4	0613	Standard	10K TPM	Succeede	7/5/2024	Defau

Step 1

Thursday, December 14, 2023 11:50 PM



File Edit Selection View Go Run ···
BonBon_Assignment [WSL: Ubuntu-22.04]
EXPLORER
BONBON_ASSIGNMENT [WSL: UBUNTU-22.04]
data
.env
.env:Zone.Identifier
assignment.ipynb
assignment.ipynb:Zone.Identifier
image.png
readme.md
requirements.txt
assignment.ipynb
readme.md
.env:Zone.Identifier
requirements.txt
.env
assignment.ipynb > M↓Init variables > import openai, os
+ Code + Markdown Run All Restart Clear All Outputs Go To Variables Outline ···
bonbon_ass (Python 3.11.5)
OUTLINE
TIMELINE
DOCKER CONTAINERS
DOCKER IMAGES
AZURE CONTAINER REGISTRY
DOCKER HUB
SUGGESTED DOCKER HUB IMAGES
PROBLEMS 13 OUTPUT DEBUG CONSOLE TERMINAL PORTS 6 JUPYTER
bash BonBo...
bash BonBo...
WSL: Ubuntu-22.04 ⊗ 7 ⚠ 6 6
Cell 12 of 19 Prettier
ENG
11:50 PM
12/14/2023

You need to set value of `OPENAI_API_KEY` that you get from the training team in the .env file

```python
import openai, os
from dotenv import load_dotenv

load_dotenv()
openai.api_type = "azure"
openai.api_version = "2023-07-01-preview"
```

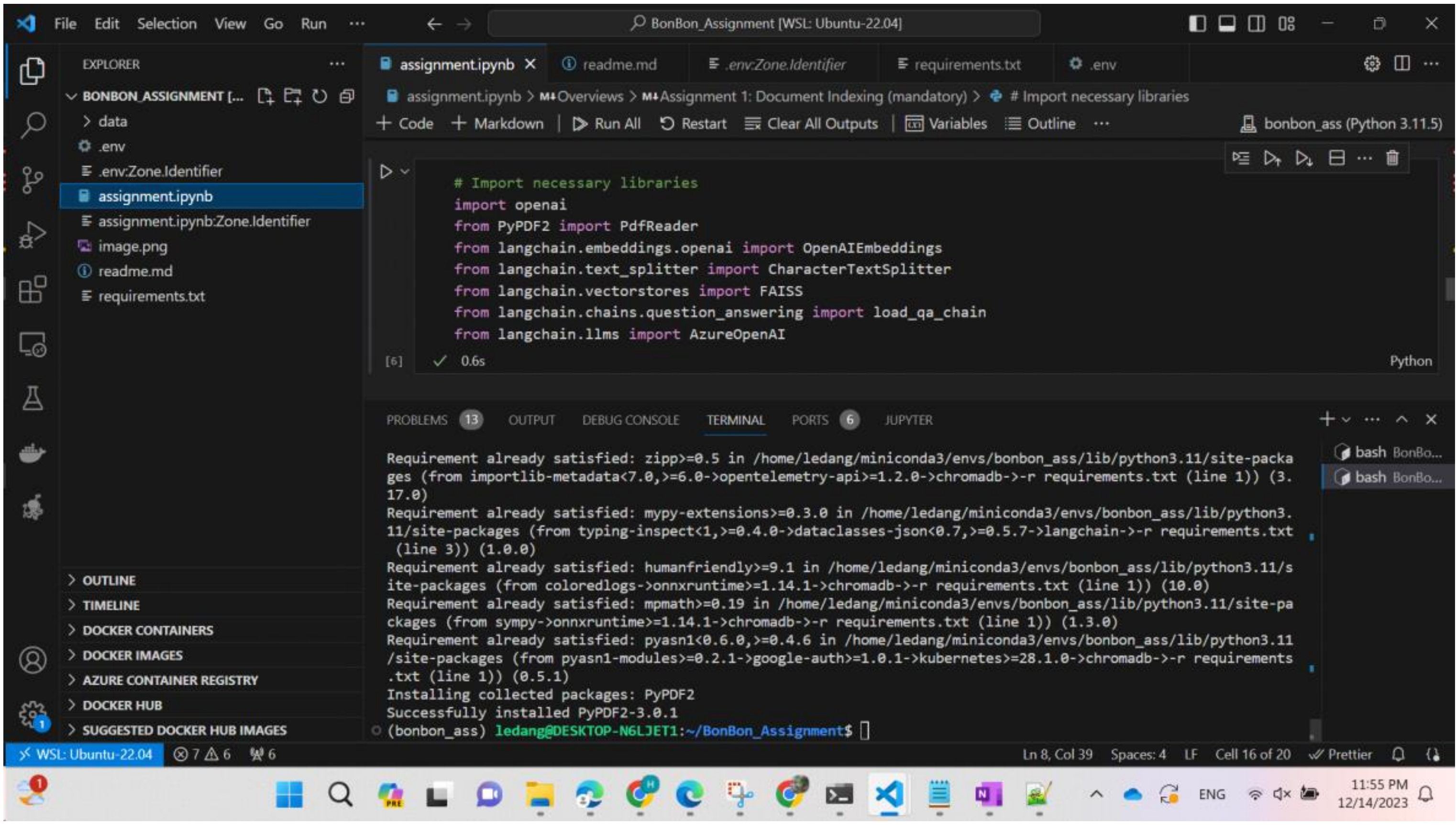
[4] ✓ 0.1s Python

Overviews

```
try-instrumentation-0.42b0 opentelemetry-instrumentation-asgi-0.42b0 opentelemetry-instrumentation-fastapi-0.
42b0 opentelemetry-proto-1.21.0 opentelemetry-sdk-1.21.0 opentelemetry-semantic-conventions-0.42b0 openteleme
try-util-http-0.42b0 overrides-7.4.0 pdfminer-20191125 plotly-5.10.0 posthog-3.1.0 protobuf-4.25.1 pulsar-cli
ent-3.3.0 pyasn1-0.5.1 pyasn1-modules-0.3.0 pycryptodome-3.19.0 pydantic-2.5.2 pydantic-core-2.14.5 pypika-0.
48.9 python-dotenv-1.0.0 requests-2.31.0 requests-oauthlib-1.3.1 rsa-4.9 sniffio-1.3.0 starlette-0.27.0 sympy
-1.12 tenacity-8.2.3 tokenizers-0.15.0 tqdm-4.66.1 typer-0.9.0 types-requests-2.31.0.6 types-urllib3-1.26.25.
14 typing-inspect-0.9.0 urllib3-1.26.18 uvicorn-0.24.0.post1 uvloop-0.19.0 watchfiles-0.21.0 websocket-client
-1.7.0 websockets-12.0 wrapt-1.16.0 yarl-1.9.4
(bonbon_ass) ledang@DESKTOP-N6LJET1:~/BonBon_Assignment$ 
```

Step 2

Thursday, December 14, 2023 11:54 PM

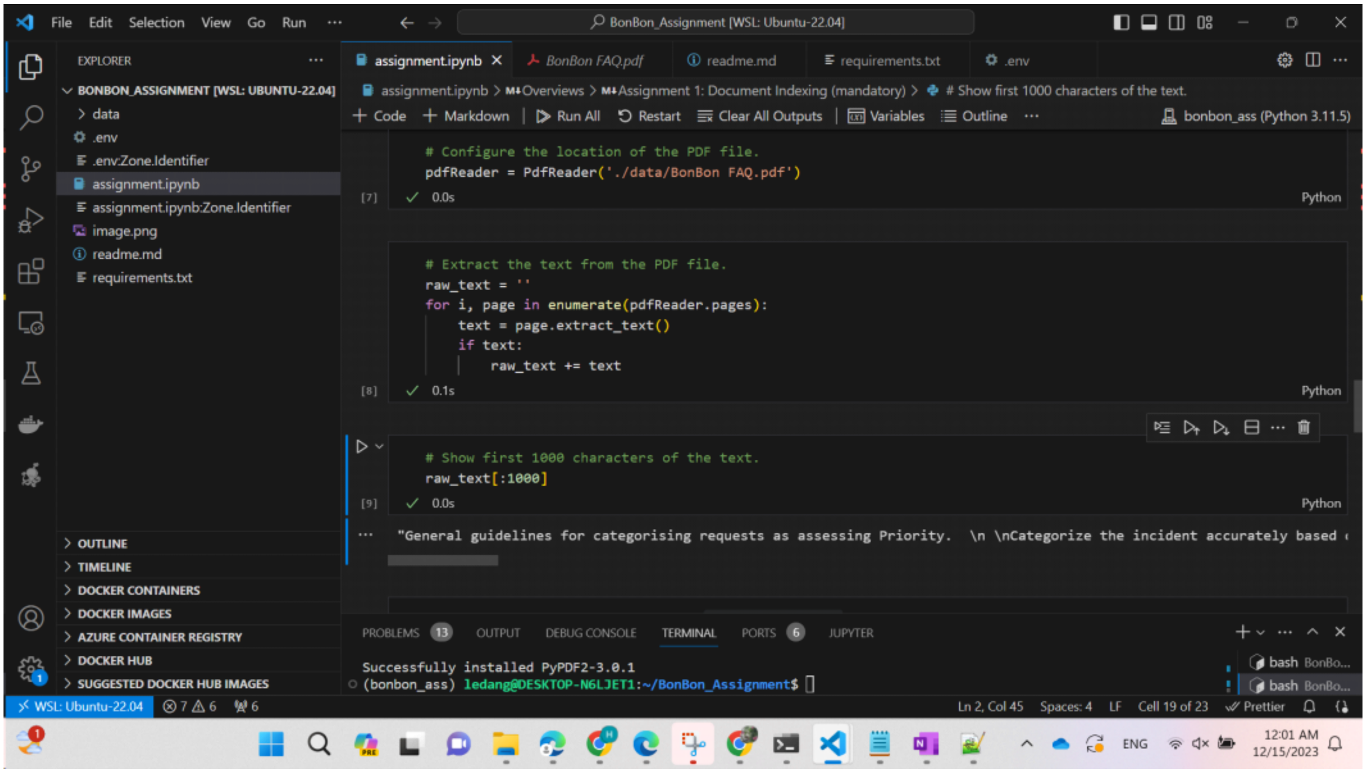


```python
# Import necessary libraries
import openai
from PyPDF2 import PdfReader
from langchain.embeddings.openai import OpenAIEmbeddings
from langchain.text_splitter import CharacterTextSplitter
from langchain.vectorstores import FAISS
from langchain.chains.question_answering import load_qa_chain
from langchain.llms import AzureOpenAI
```

```
Requirement already satisfied: zipp>=0.5 in /home/ledang/miniconda3/envs/bonbon_ass/lib/python3.11/site-packa
ges (from importlib-metadata<7.0,>=6.0->opentelemetry-api>=1.2.0->chromadb->-r requirements.txt (line 1)) (3.
17.0)
Requirement already satisfied: mypy-extensions>=0.3.0 in /home/ledang/miniconda3/envs/bonbon_ass/lib/python3.
11/site-packages (from typing-inspect<1,>=0.4.0->dataclasses-json<0.7,>=0.5.7->langchain->-r requirements.txt
 (line 3)) (1.0.0)
Requirement already satisfied: humanfriendly>=9.1 in /home/ledang/miniconda3/envs/bonbon_ass/lib/python3.11/s
ite-packages (from coloredlogs->onnxruntime>=1.14.1->chromadb->-r requirements.txt (line 1)) (10.0)
Requirement already satisfied: mpmath>=0.19 in /home/ledang/miniconda3/envs/bonbon_ass/lib/python3.11/site-pa
ckages (from sympy->onnxruntime>=1.14.1->chromadb->-r requirements.txt (line 1)) (1.3.0)
Requirement already satisfied: pyasn1<0.6.0,>=0.4.6 in /home/ledang/miniconda3/envs/bonbon_ass/lib/python3.11
/site-packages (from pyasn1-modules>=0.2.1->google-auth>=1.0.1->kubernetes>=28.1.0->chromadb->-r requirements
.txt (line 1)) (0.5.1)
Installing collected packages: PyPDF2
Successfully installed PyPDF2-3.0.1
(bonbon_ass) ledang@DESKTOP-N6LJET1:~/BonBon_Assignment$ 
```

Step 3

Thursday, December 14, 2023 11:55 PM



```
# Configure the location of the PDF file.
pdfReader = PdfReader('./data/BonBon FAQ.pdf')
```

[7] ✓ 0.0s

```
# Extract the text from the PDF file.
raw_text = ''
for i, page in enumerate(pdfReader.pages):
    text = page.extract_text()
    if text:
        raw_text += text
```

[8] ✓ 0.1s

```
# Show first 1000 characters of the text.
raw_text[:1000]
```

[9] ✓ 0.0s

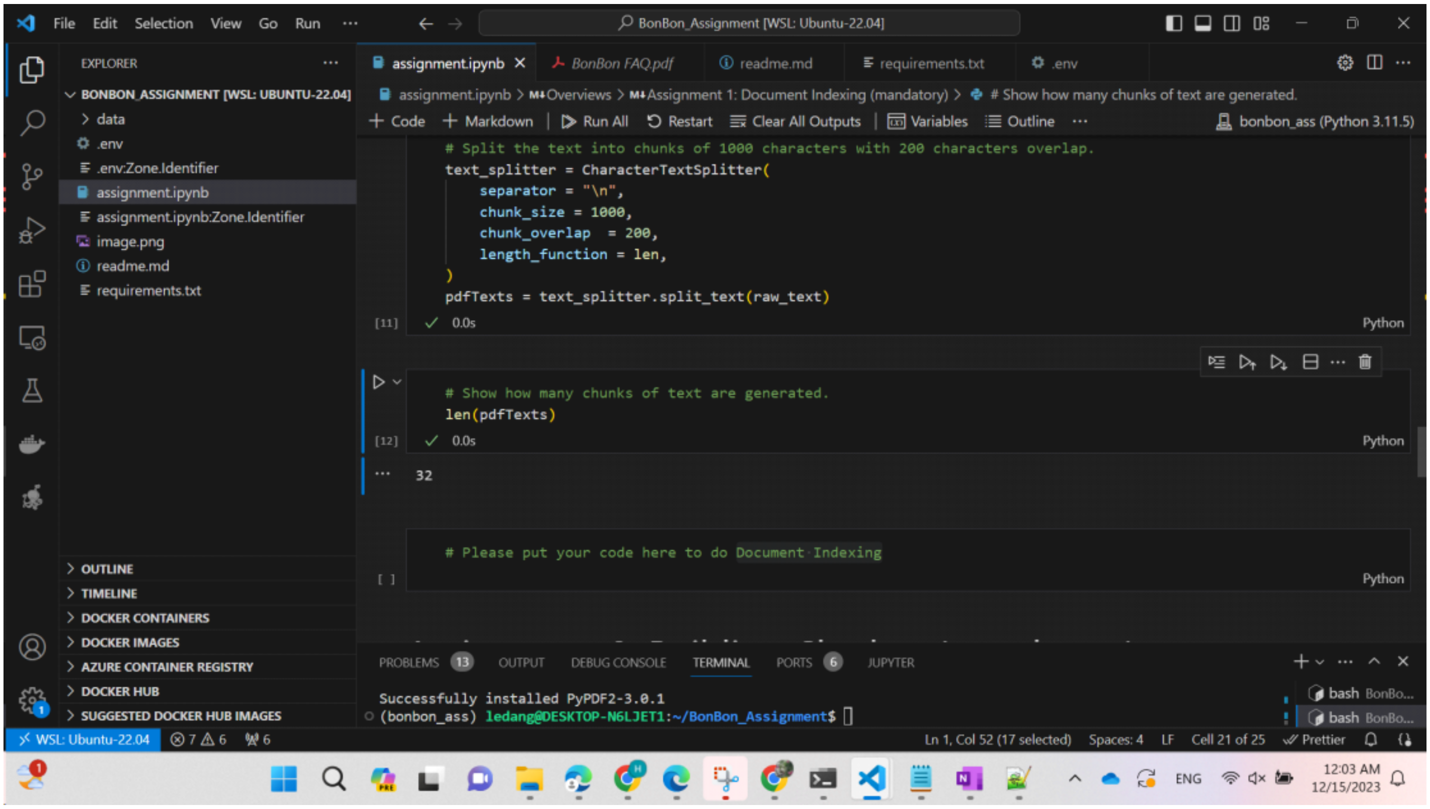
```
"General guidelines for categorising requests as assessing Priority.  \n \nCategorize the incident accurately based
```

```
Successfully installed PyPDF2-3.0.1
(bonbon_ass) ledang@DESKTOP-N6LJET1:~/BonBon_Assignment$
```

Reference:
https://github.com/easonlai/azure_openai_langchain_sample/blob/main/chat_with_pdf.ipynb

Step 4

Friday, December 15, 2023 12:02 AM


File Edit Selection View Go Run ...
BonBon_Assignment [WSL: Ubuntu-22.04]
EXPLORER
BONBON_ASSIGNMENT [WSL: UBUNTU-22.04]
data
.env
.env:Zone.Identifier
assignment.ipynb
assignment.ipynb:Zone.Identifier
image.png
readme.md
requirements.txt
OUTLINE
TIMELINE
DOCKER CONTAINERS
DOCKER IMAGES
AZURE CONTAINER REGISTRY
DOCKER HUB
SUGGESTED DOCKER HUB IMAGES
assignment.ipynb
BonBon FAQ.pdf
readme.md
requirements.txt
.env
assignment.ipynb > Overviews > Assignment 1: Document Indexing (mandatory) > # Show how many chunks of text are generated.
+ Code + Markdown | Run All Restart Clear All Outputs | Variables Outline ...
bonbon_ass (Python 3.11.5)
PROBLEMS 13 OUTPUT DEBUG CONSOLE TERMINAL PORTS 6 JUPYTER
bash BonBo...
bash BonBo...
WSL: Ubuntu-22.04 ⊗ 7 ⚠ 6 6
Ln 1, Col 52 (17 selected) Spaces: 4 LF Cell 21 of 25 Prettier
ENG
12:03 AM
12/15/2023

```
# Split the text into chunks of 1000 characters with 200 characters overlap.
text_splitter = CharacterTextSplitter(
    separator = "\n",
    chunk_size = 1000,
    chunk_overlap  = 200,
    length_function = len,
)
pdfTexts = text_splitter.split_text(raw_text)
```

[11] ✓ 0.0s Python

```
# Show how many chunks of text are generated.
len(pdfTexts)
```

[12] ✓ 0.0s Python

```
32
```

```
# Please put your code here to do Document Indexing
```

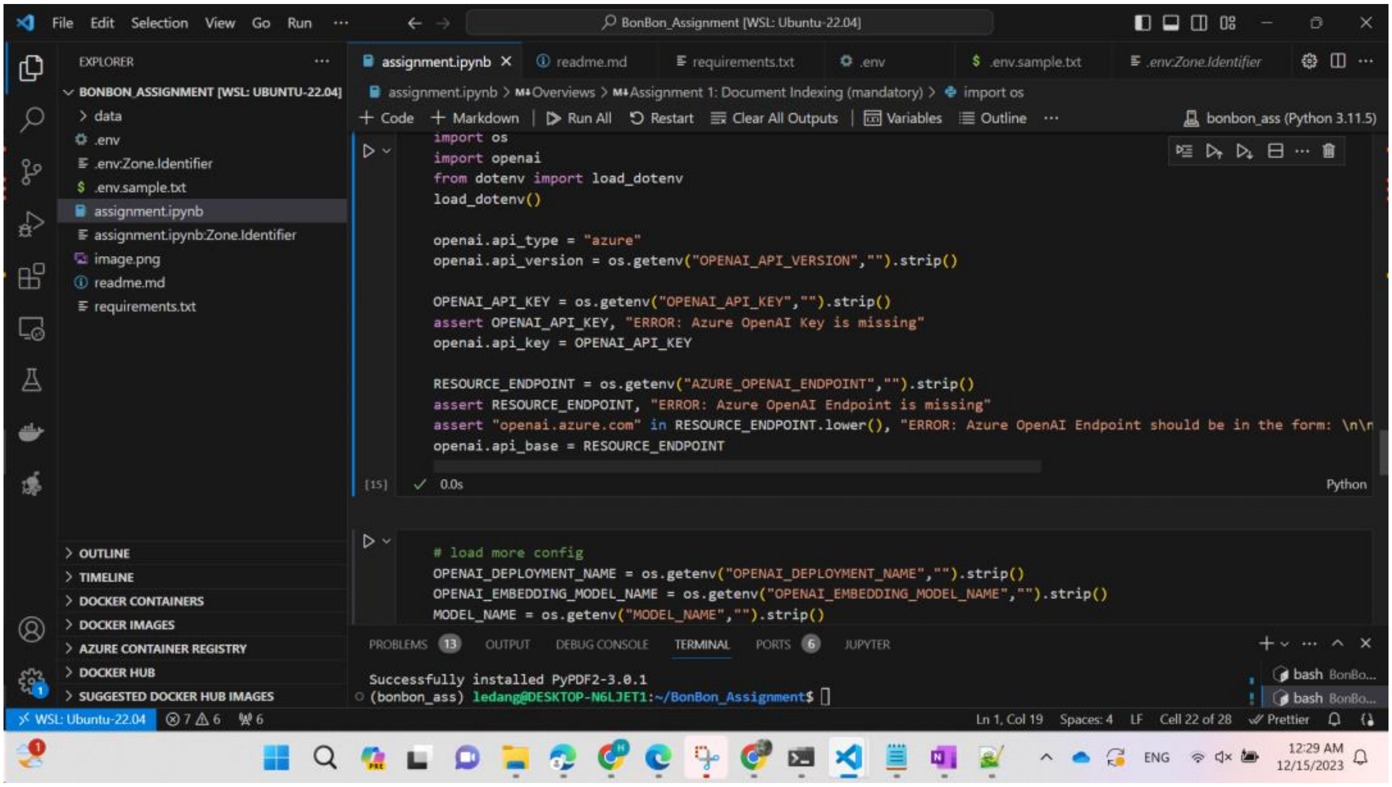
[ ] Python

```
Successfully installed PyPDF2-3.0.1
(bonbon_ass) ledang@DESKTOP-N6LJET1:~/BonBon_Assignment$ 
```

Reference: https://github.com/phaledang/clone-azure-openai-samples/blob/main/quick_start/02_OpenAI_getting_started.ipynb

Step 5: Load config, init embeding

Friday, December 15, 2023 12:27 AM

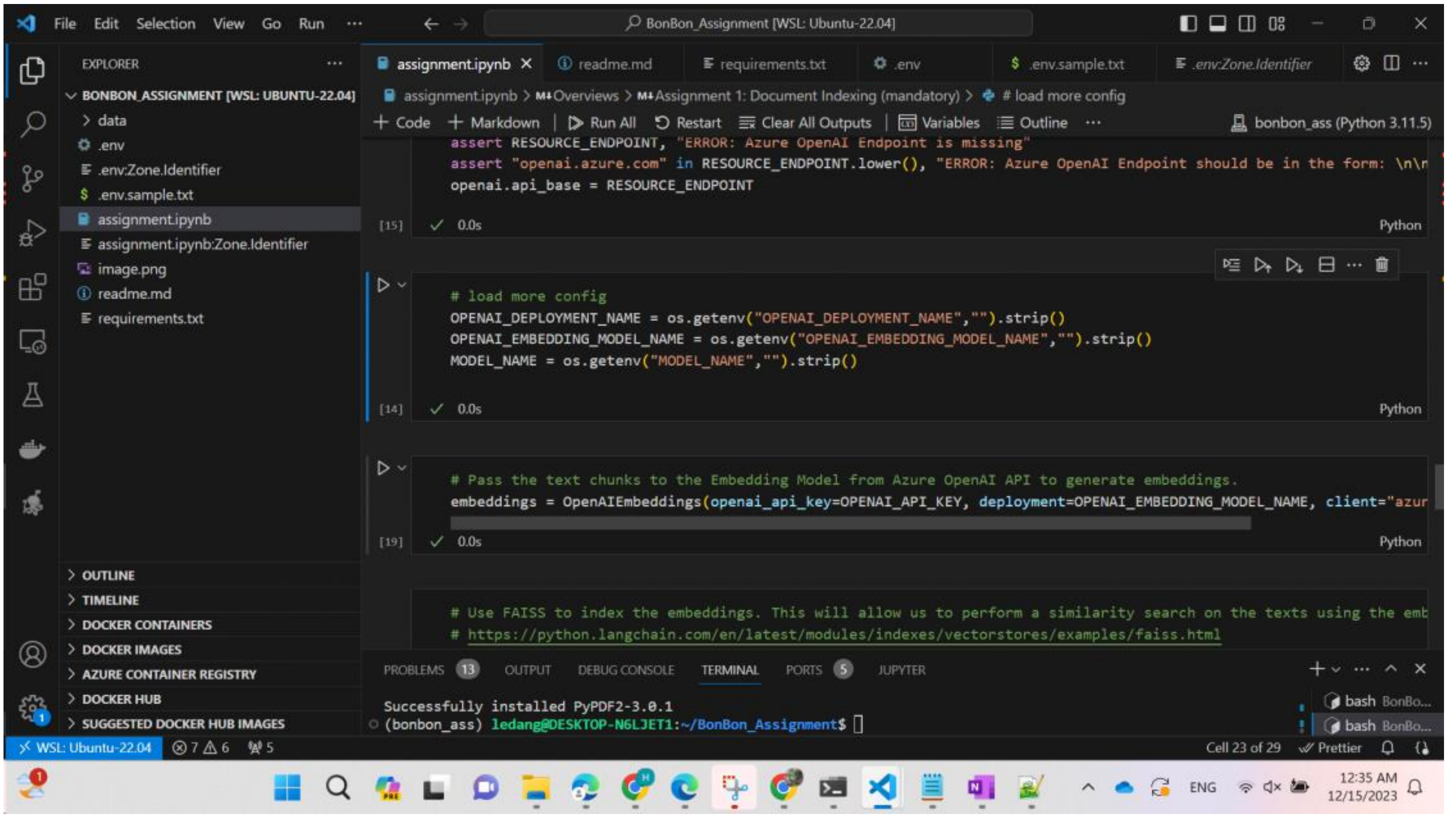

BonBon_Assignment [WSL: Ubuntu-22.04]
assignment.ipynb
readme.md
requirements.txt
.env
.env.sample.txt
.env:Zone.Identifier
assignment.ipynb > Overviews > Assignment 1: Document Indexing (mandatory) > import os
bonbon_ass (Python 3.11.5)
EXPLORER
BONBON_ASSIGNMENT [WSL: UBUNTU-22.04]
data
assignment.ipynb:Zone.Identifier
image.png
PROBLEMS 13
TERMINAL
PORTS 6
WSL: Ubuntu-22.04
Ln 1, Col 19
Cell 22 of 28
12:29 AM
12/15/2023

```python
import os
import openai
from dotenv import load_dotenv
load_dotenv()

openai.api_type = "azure"
openai.api_version = os.getenv("OPENAI_API_VERSION","").strip()

OPENAI_API_KEY = os.getenv("OPENAI_API_KEY","").strip()
assert OPENAI_API_KEY, "ERROR: Azure OpenAI Key is missing"
openai.api_key = OPENAI_API_KEY

RESOURCE_ENDPOINT = os.getenv("AZURE_OPENAI_ENDPOINT","").strip()
assert RESOURCE_ENDPOINT, "ERROR: Azure OpenAI Endpoint is missing"
assert "openai.azure.com" in RESOURCE_ENDPOINT.lower(), "ERROR: Azure OpenAI Endpoint should be in the form: \n\n
openai.api_base = RESOURCE_ENDPOINT
```

[15] ✓ 0.0s

```python
# load more config
OPENAI_DEPLOYMENT_NAME = os.getenv("OPENAI_DEPLOYMENT_NAME","").strip()
OPENAI_EMBEDDING_MODEL_NAME = os.getenv("OPENAI_EMBEDDING_MODEL_NAME","").strip()
MODEL_NAME = os.getenv("MODEL_NAME","").strip()
```

```
Successfully installed PyPDF2-3.0.1
(bonbon_ass) ledang@DESKTOP-N6LJET1:~/BonBon_Assignment$
```


BonBon_Assignment [WSL: Ubuntu-22.04]
assignment.ipynb > Overviews > Assignment 1: Document Indexing (mandatory) > # load more config
bonbon_ass (Python 3.11.5)
PROBLEMS 13
TERMINAL
PORTS 5
WSL: Ubuntu-22.04
Cell 23 of 29
12:35 AM
12/15/2023

```python
assert RESOURCE_ENDPOINT, "ERROR: Azure OpenAI Endpoint is missing"
assert "openai.azure.com" in RESOURCE_ENDPOINT.lower(), "ERROR: Azure OpenAI Endpoint should be in the form: \n\r
openai.api_base = RESOURCE_ENDPOINT
```

[15] ✓ 0.0s

```python
# load more config
OPENAI_DEPLOYMENT_NAME = os.getenv("OPENAI_DEPLOYMENT_NAME","").strip()
OPENAI_EMBEDDING_MODEL_NAME = os.getenv("OPENAI_EMBEDDING_MODEL_NAME","").strip()
MODEL_NAME = os.getenv("MODEL_NAME","").strip()
```

[14] ✓ 0.0s

```python
# Pass the text chunks to the Embedding Model from Azure OpenAI API to generate embeddings.
embeddings = OpenAIEmbeddings(openai_api_key=OPENAI_API_KEY, deployment=OPENAI_EMBEDDING_MODEL_NAME, client="azur
```

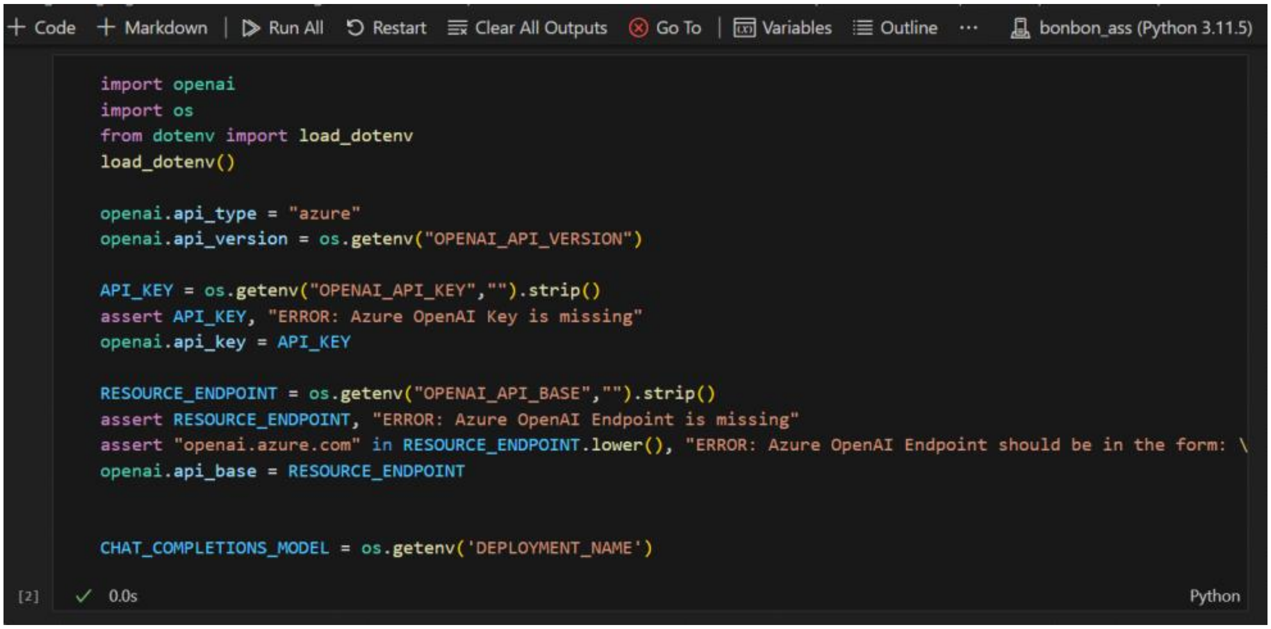
[19] ✓ 0.0s

```python
# Use FAISS to index the embeddings. This will allow us to perform a similarity search on the texts using the emb
# https://python.langchain.com/en/latest/modules/indexes/vectorstores/examples/faiss.html
```

```
Successfully installed PyPDF2-3.0.1
(bonbon_ass) ledang@DESKTOP-N6LJET1:~/BonBon_Assignment$
```

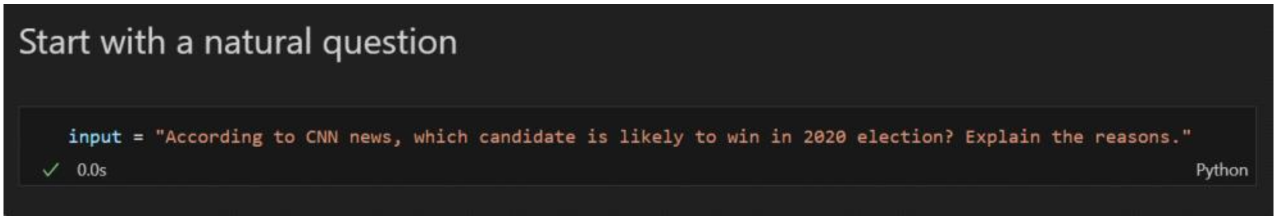
Chat

Friday, December 15, 2023



```python
import openai
import os
from dotenv import load_dotenv
load_dotenv()

openai.api_type = "azure"
openai.api_version = os.getenv("OPENAI_API_VERSION")

API_KEY = os.getenv("OPENAI_API_KEY","").strip()
assert API_KEY, "ERROR: Azure OpenAI Key is missing"
openai.api_key = API_KEY

RESOURCE_ENDPOINT = os.getenv("OPENAI_API_BASE","").strip()
assert RESOURCE_ENDPOINT, "ERROR: Azure OpenAI Endpoint is missing"
assert "openai.azure.com" in RESOURCE_ENDPOINT.lower(), "ERROR: Azure OpenAI Endpoint should be in the form: \
openai.api_base = RESOURCE_ENDPOINT


CHAT_COMPLETIONS_MODEL = os.getenv('DEPLOYMENT_NAME')
```

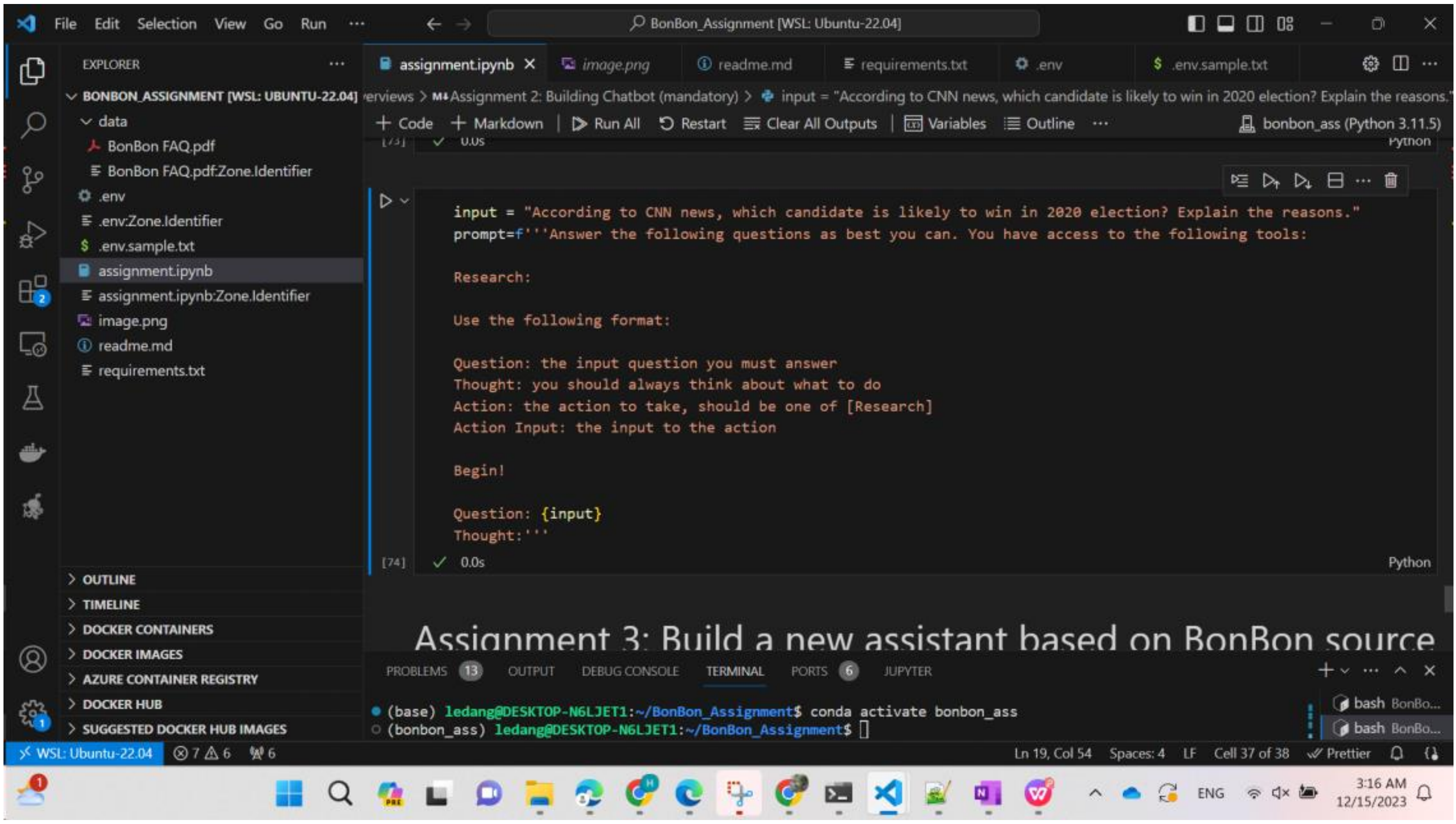
[2] ✓ 0.0s Python



Start with a natural question

```python
input = "According to CNN news, which candidate is likely to win in 2020 election? Explain the reasons."
```

✓ 0.0s Python



```python
input = "According to CNN news, which candidate is likely to win in 2020 election? Explain the reasons."
prompt=f'''Answer the following questions as best you can. You have access to the following tools:

Research:

Use the following format:

Question: the input question you must answer
Thought: you should always think about what to do
Action: the action to take, should be one of [Research]
Action Input: the input to the action

Begin!

Question: {input}
Thought:'''
```

[74] ✓ 0.0s Python

Assignment 3: Build a new assistant based on BonBon source

```
(base) ledang@DESKTOP-N6LJET1:~/BonBon_Assignment$ conda activate bonbon_ass
(bonbon_ass) ledang@DESKTOP-N6LJET1:~/BonBon_Assignment$ 
```

Friday, December 15, 2023 9:11 PM

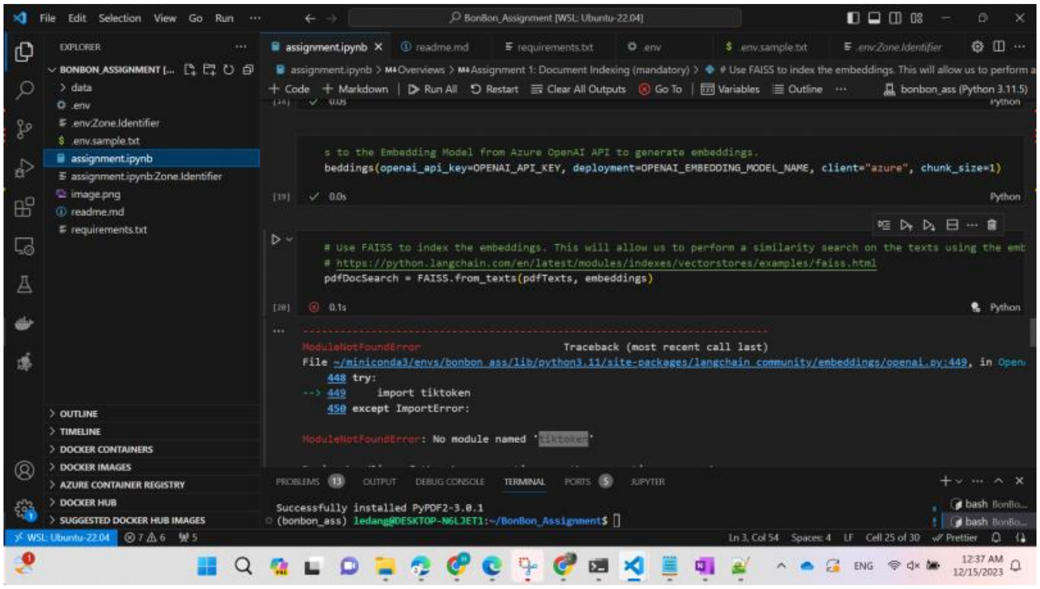
```
## Issues:
###  Command 'python' not found
Edit file .bashrc to add
alias python='python3'
```

Run command

```
sudo apt-get install python-is-python3
```

Fix Issue no tiktoken

Friday, December 15, 2023 12:35 AM



```
s to the Embedding Model from Azure OpenAI API to generate embeddings.
beddings(openai_api_key=OPENAI_API_KEY, deployment=OPENAI_EMBEDDING_MODEL_NAME, client="azure", chunk_size=1)
```

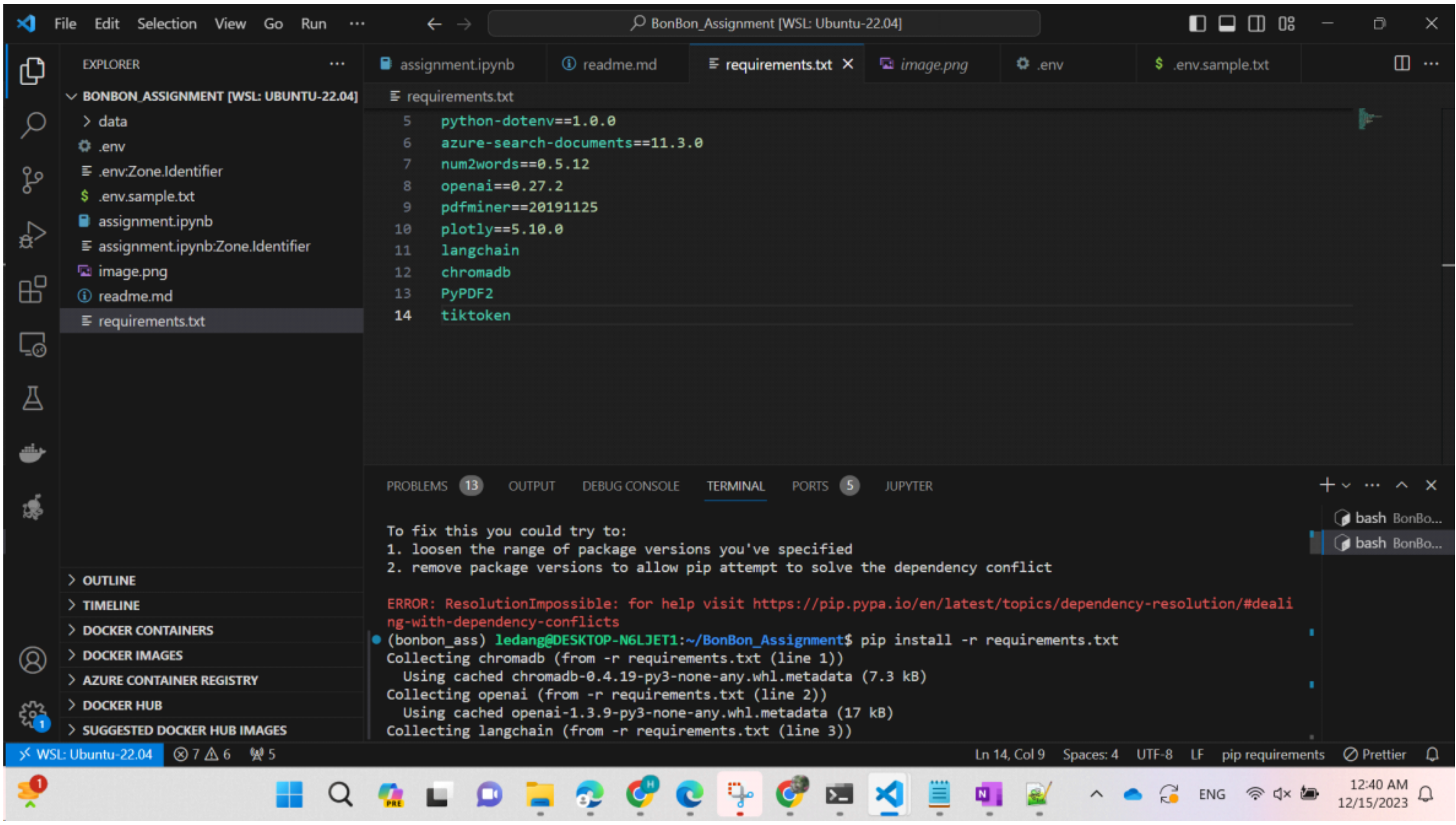
```
# Use FAISS to index the embeddings. This will allow us to perform a similarity search on the texts using the emb
# https://python.langchain.com/en/latest/modules/indexes/vectorstores/examples/faiss.html
pdfDocSearch = FAISS.from_texts(pdfTexts, embeddings)
```

```
---------------------------------------------------------------------------
ModuleNotFoundError                       Traceback (most recent call last)
File ~/miniconda3/envs/bonbon_ass/lib/python3.11/site-packages/langchain_community/embeddings/openai.py:449, in Open
    448 try:
--> 449     import tiktoken
    450 except ImportError:

ModuleNotFoundError: No module named 'tiktoken'
```

```
Successfully installed PyPDF2-3.0.1
(bonbon_ass) ledang@DESKTOP-N6LJET1:~/BonBon_Assignment$
```

Update requirement.txt and run

pip install -r requirements.txt



```
python-dotenv==1.0.0
azure-search-documents==11.3.0
num2words==0.5.12
openai==0.27.2
pdfminer==20191125
plotly==5.10.0
langchain
chromadb
PyPDF2
tiktoken
```

```
To fix this you could try to:
1. loosen the range of package versions you've specified
2. remove package versions to allow pip attempt to solve the dependency conflict

ERROR: ResolutionImpossible: for help visit https://pip.pypa.io/en/latest/topics/dependency-resolution/#dealing-with-dependency-conflicts
(bonbon_ass) ledang@DESKTOP-N6LJET1:~/BonBon_Assignment$ pip install -r requirements.txt
Collecting chromadb (from -r requirements.txt (line 1))
  Using cached chromadb-0.4.19-py3-none-any.whl.metadata (7.3 kB)
Collecting openai (from -r requirements.txt (line 2))
  Using cached openai-1.3.9-py3-none-any.whl.metadata (17 kB)
Collecting langchain (from -r requirements.txt (line 3))
```

References

Friday, December 15, 2023 12:15 AM

https://github.com/phaledang/clone-azure-openai-samples/blob/main/quick_start/02_OpenAI_getting_started.ipynb

https://python.langchain.com/docs/integrations/vectorstores/faiss

https://www.reddit.com/r/LangChain/comments/146y8bd/debugging_faissfrom_textsdocs_embeddings/

https://github.com/phaledang/ai-python/wiki

Git sample
https://github.com/easonlai/azure_openai_langchain_sample/blob/main/README.md
https://github.com/Azure/azure-openai-samples/blob/main/quick_start/08_prompt_engineering.ipynb
https://github.com/hwchase17/chroma-langchain