

Date: 17 October 2023

Team ID: 329 PROJECT

ID: Proj_227277_Team_1

NAME: Dilipvijayakumar T

Installing Packages

1. Package name: transformers

Use: For GPT-3 integration

Command to install: pip install transformers

```
Command Prompt
----- 61.5/61.5 kB 649.3 kB/s eta 0:00:00
Collecting urllib3<3,>=1.21.1 (from requests->transformers)
  Downloading urllib3-2.0.6-py3-none-any.whl (123 kB)
----- 123.0/123.0 kB 908.5 kB/s eta 0:00:00
Collecting certifi>=2017.4.17 (from requests->transformers)
  Downloading certifi-2023.7.22-py3-none-any.whl (158 kB)
----- 158.3/158.3 kB 231.4 kB/s eta 0:00:00
Installing collected packages: urllib3, typing-extensions, safetensors, regex, pyyaml, packaging, numpy, idna, fsspec, filelock, colorama, charset-normalizer, certifi,
tqdm, requests, huggingface-hub, tokenizers, transformers
Successfully installed certifi-2023.7.22 charset-normalizer-3.3.0 colorama-0.4.6 filelock-3.12.4 fsspec-2023.9.2 huggingface-hub-0.17.3 idna-3.4 numpy-1.26.1 packaging-
23.2 pyyaml-6.0.1 regex-2023.10.3 requests-2.31.0 safetensors-0.4.0 tokenizers-0.14.1 tqdm-4.66.1 transformers-4.34.0 typing-extensions-4.8.0 urllib3-2.0.6

[notice] A new release of pip is available: 23.1.2 -> 23.3
[notice] To update, run: python.exe -m pip install --upgrade pip

C:\Users\Anbarasan>
C:\Users\Anbarasan>pip install Flask
Collecting Flask
  Downloading flask-3.0.0-py3-none-any.whl (99 kB)
----- 99.7/99.7 kB 520.3 kB/s eta 0:00:00
Collecting Werkzeug>=3.0.0 (from Flask)
  Downloading werkzeug-3.0.0-py3-none-any.whl (226 kB)
----- 226.6/226.6 kB 1.1 MB/s eta 0:00:00
Collecting Jinja2>=3.1.2 (from Flask)
  Downloading Jinja2-3.1.2-py3-none-any.whl (133 kB)
----- 133.1/133.1 kB 1.1 MB/s eta 0:00:00
Collecting itsdangerous>=2.1.2 (from Flask)
  Downloading itsdangerous-2.1.2-py3-none-any.whl (15 kB)
Collecting click>=8.1.3 (from Flask)
  Downloading click-8.1.7-py3-none-any.whl (97 kB)
----- 97.9/97.9 kB 622.0 kB/s eta 0:00:00
Collecting blinker>=1.6.2 (from Flask)
  Downloading blinker-1.6.3-py3-none-any.whl (13 kB)
Requirement already satisfied: colorama in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (from click>=8.1.3->Flask) (0.4.6)
Collecting MarkupSafe>=2.0 (from Jinja2>=3.1.2->Flask)
  Downloading MarkupSafe-2.1.3-cp311-cp311-win_amd64.whl (17 kB)
Installing collected packages: MarkupSafe, itsdangerous, click, blinker, Werkzeug, Jinja2, Flask
Successfully installed Flask-3.0.0 Jinja2-3.1.2 MarkupSafe-2.1.3 Werkzeug-3.0.0 blinker-1.6.3 click-8.1.7 itsdangerous-2.1.2

[notice] A new release of pip is available: 23.1.2 -> 23.3
[notice] To update, run: python.exe -m pip install --upgrade pip

C:\Users\Anbarasan>
```

2. Package name: Flask
Use: For web app development
Command to install: pip install Flask

```
Command Prompt
[notice] A new release of pip is available: 23.1.2 -> 23.3
[notice] To update, run: python.exe -m pip install --upgrade pip

C:\Users\Anbarasan>pip install transformers
Requirement already satisfied: transformers in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (4.34.0)
Requirement already satisfied: filelock in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (from transformers) (3.12.4)
Requirement already satisfied: huggingface-hub<1.0,>=0.16.4 in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (from transformers) (0.17.3)
Requirement already satisfied: numpy>=1.17 in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (from transformers) (1.26.1)
Requirement already satisfied: packaging>=20.0 in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (from transformers) (23.2)
Requirement already satisfied: pyyaml>=5.1 in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (from transformers) (6.0.1)
Requirement already satisfied: regex<=2019.12.17 in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (from transformers) (2023.10.3)
Requirement already satisfied: requests in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (from transformers) (2.31.0)
Requirement already satisfied: tokenizers<0.15,>=0.14 in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (from transformers) (0.14.1)
Requirement already satisfied: safetensors>=0.3.1 in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (from transformers) (0.4.0)
Requirement already satisfied: tqdm>=4.27 in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (from transformers) (4.66.1)
Requirement already satisfied: fsspec in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (from huggingface-hub<1.0,>=0.16.4->transformers) (2023.9.2)
Requirement already satisfied: typing-extensions>=3.7.4.3 in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (from huggingface-hub<1.0,>=0.16.4->transformers) (4.8.0)
Requirement already satisfied: colorama in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (from tqdm>=4.27->transformers) (0.4.6)
Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (from requests->transformers) (3.3.0)
Requirement already satisfied: idna<4,>=2.5 in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (from requests->transformers) (3.4)
Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (from requests->transformers) (2.0.6)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\anbarasan\appdata\local\programs\python\python311\lib\site-packages (from requests->transformers) (2023.7.22)

[notice] A new release of pip is available: 23.1.2 -> 23.3
[notice] To update, run: python.exe -m pip install --upgrade pip

C:\Users\Anbarasan>
```

Program for basic chat bot conversation

I provided source code file called “AI_Phase3_source_code.ipynb” in my git hub repository

```
# import all required libraries
import numpy as np
import string
from nltk.corpus import stopwords
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.neural_network import MLPClassifier
from sklearn.feature_extraction.text import TfidfTransformer, TfidfVectorizer
from sklearn.pipeline import Pipeline
```

```
# importing the dataset
df = pd.read_csv(r"C:\Users\COMPAQ\Desktop\IBM\dataset\dialogs.txt", sep='\t')

df.head()
```

	hi, how are you doing?	i'm fine. how about yourself?
0	i'm fine. how about yourself?	i'm pretty good. thanks for asking.
1	i'm pretty good. thanks for asking.	no problem. so how have you been?
2	no problem. so how have you been?	i've been great. what about you?
3	i've been great. what about you?	i've been good. i'm in school right now.
4	i've been good. i'm in school right now.	what school do you go to?

#add column names

```
df.columns=['Questions','Answers']
df
```

	Questions	Answers
0	i'm fine. how about yourself?	i'm pretty good. thanks for asking.
1	i'm pretty good. thanks for asking.	no problem. so how have you been?
2	no problem. so how have you been?	i've been great. what about you?
3	i've been great. what about you?	i've been good. i'm in school right now.
4	i've been good. i'm in school right now.	what school do you go to?
...
3719	that's a good question. maybe it's not old age.	are you right-handed?
3720	are you right-handed?	yes. all my life.
3721	yes. all my life.	you're wearing out your right hand. stop using...
3722	you're wearing out your right hand. stop using...	but i do all my writing with my right hand.
3723	but i do all my writing with my right hand.	start typing instead. that way your left hand ...
3724 rows × 2 columns		

Data Preprocessing

```
# Assuming you have a DataFrame 'df' with 'Questions' and 'Answers' columns

df['question tokens'] = df['Questions'].apply(lambda x: len(x.split()))
df['answer tokens'] = df['Answers'].apply(lambda x: len(x.split()))

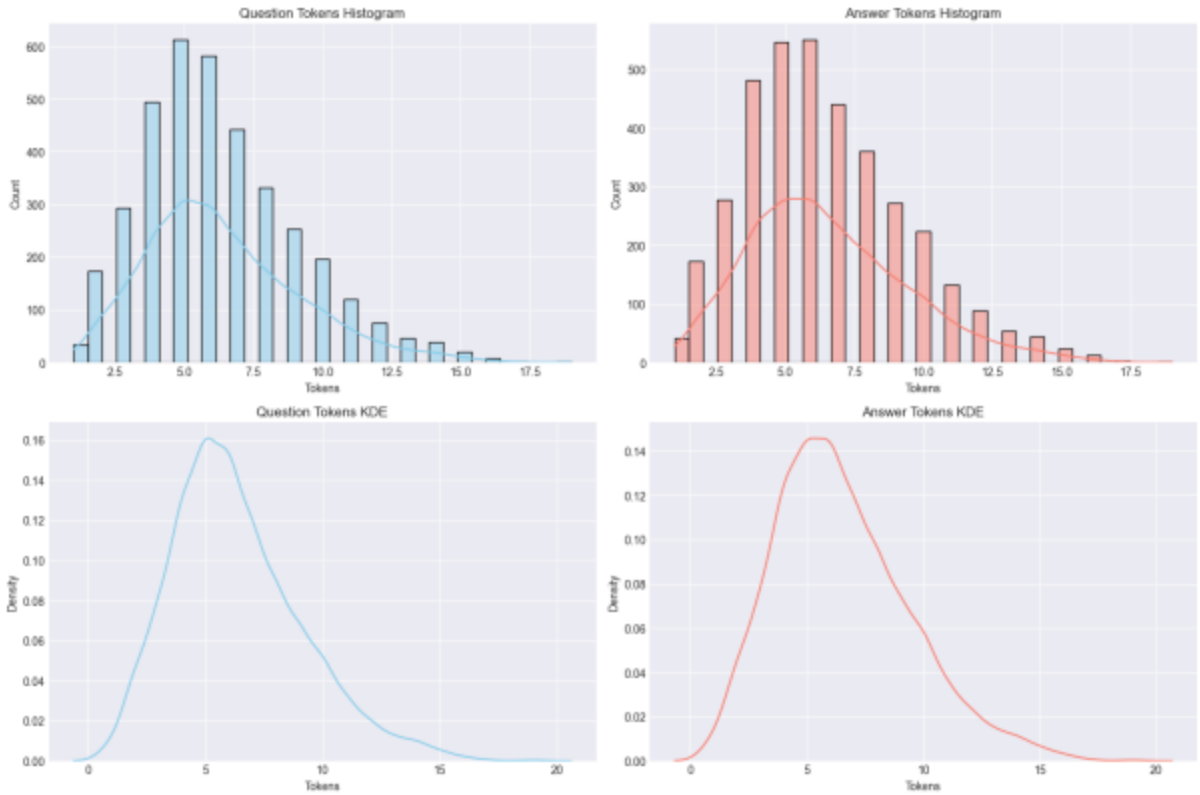
plt.style.use('fivethirtyeight')
fig, ax = plt.subplots(nrows=1, ncols=3, figsize=(20, 5))
sns.set_palette('Set2')

# Create bar plots for question tokens and answer tokens
sns.barplot(x='question tokens', y=df.index, data=df, ax=ax[0])
ax[0].set_xlabel('Question Tokens')
ax[0].set_ylabel('Index')
ax[0].set_title('Question Tokens Bar Plot')

sns.barplot(x='answer tokens', y=df.index, data=df, ax=ax[1])
ax[1].set_xlabel('Answer Tokens')
ax[1].set_ylabel('Index')
ax[1].set_title('Answer Tokens Bar Plot')

# Create a scatter plot with a regression line for the relationship between question and
answer tokens
sns.regplot(x='question tokens', y='answer tokens', data=df, ax=ax[2],
scatter_kws={'alpha':0.5})
ax[2].set_xlabel('Question Tokens')
ax[2].set_ylabel('Answer Tokens')
ax[2].set_title('Scatter Plot of Question Tokens vs. Answer Tokens')

plt.tight_layout()
plt.show()
```



#Function for converting upper to lower case

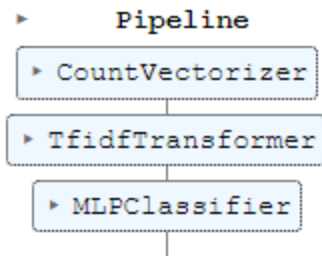
def cleaner(x):

 return [a for a in (''.join([a for a in x if a not in string.punctuation])).lower().split()]

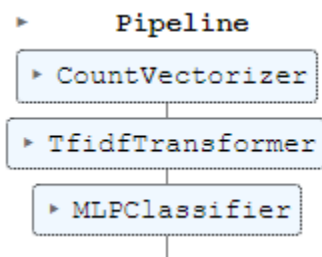
#Model

```
Pipe = Pipeline([
    ('bow',CountVectorizer(analyzer=cleaner)),
    ('tfidf',TfidfTransformer()),
    ('classifier',MLPClassifier())
])
```

```
Pipe.fit(df['Questions'],df['Answers'])
```



```
Pipe.fit(df['Questions'],df['Answers'])
```



```
#Text
```

```
Pipe.predict(['like how clear the sky gets after it rains.'])[0]
```

```
'i feel the same way. it smells so good after it rains.'
```

```
Pipe.predict(['i want this trip to be perfect, i hope it stays warm.'])[0]
```

```
"this california weather is so uncertain, it's impossible to know what'll happen."
```

```
Pipe.predict(['it would not be good if it got cold this weekend.'])[0]
```

```
'i want this trip to be perfect, i hope it stays warm.'
```

```
Pipe.predict(['it would be nice if the weather would never change.'])[0]
```

```
'that would be great, then we could plan things sooner.'
```

```
Pipe.predict(['why is that?'])[0]
```

```
'because i love the snow.'
```

```
Pipe.predict(['What are you doing'])[0]
```

```
"i'm going to change the light bulb. it burnt out."
```