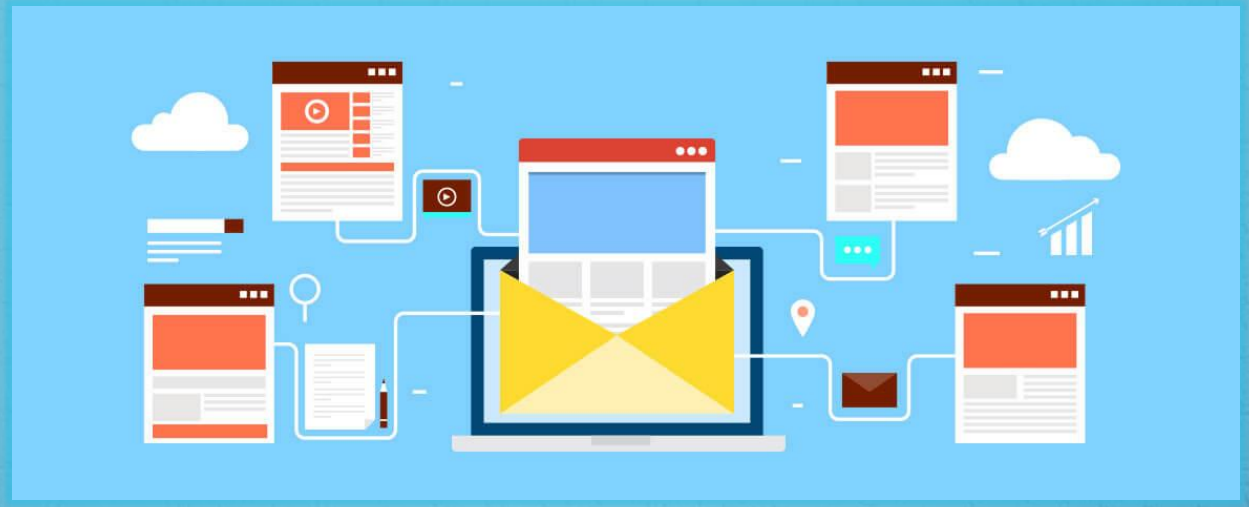


EMAIL TEMPLATE GENERATION USING NLP

KIRAN B, TEJASWINI HALMARE, SAMEER SINGH, KANCHAN KACHOT, PRATIKKUMAR TRIVEDI, APOORVA KHARWADE

BUSINESS OBJECTIVE

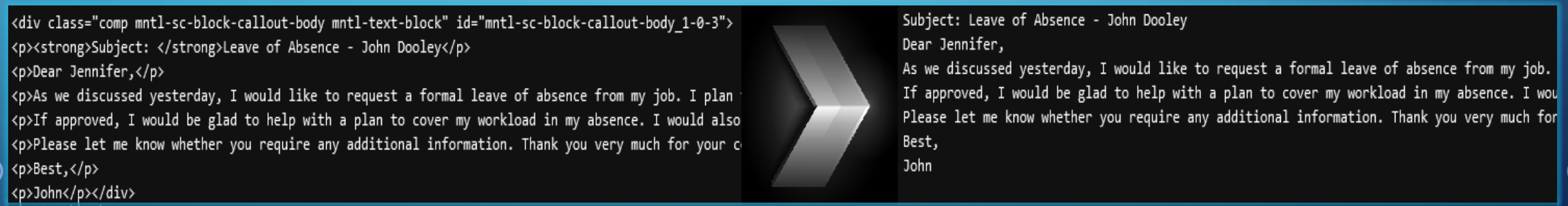
- To get the nearest email template based on the key input specified by the user
- Data(Email Samples) procurement to be done by Web-Scraping based on the user input
- Apply the necessary NLP techniques on the scraped data, to generate the template in the respective category

DATA PROCUREMENT

- Procurement of the data is done via web-scraping
- The web page has to be inspected for its structure before we begin the data extraction
- Web-scraped page contains a small section of the relevant data(email sample) and this part can be retrieved using the *div id, classname or CSS selectors*

DATA PRE-PROCESSING

- The extracted section of the data contains a lot of unwanted information such as html tags and we remove them by using the built-in python function *get_text()*
- Also, certain parts of the extracted section isn't formatted properly. This is done prior to applying the NLP techniques



EXPLORE & VISUALIZE (EDA)

- Data visualization plays a quintessential role in obtaining insights from the extracted data and also helps us to make the machine interpret the human language better
- Using Natural Language Processing techniques(NLP), we tokenize the data, remove the Stop words, identify the Parts of Speech(POS), visualize the parse tree(dependency tree), Named Entity Recognition(NER), extract the Nouns and Verbs and list their word count(which in turn is provided to the user to select and generate a respective template under the chosen email category)

EDA (CONT.…)

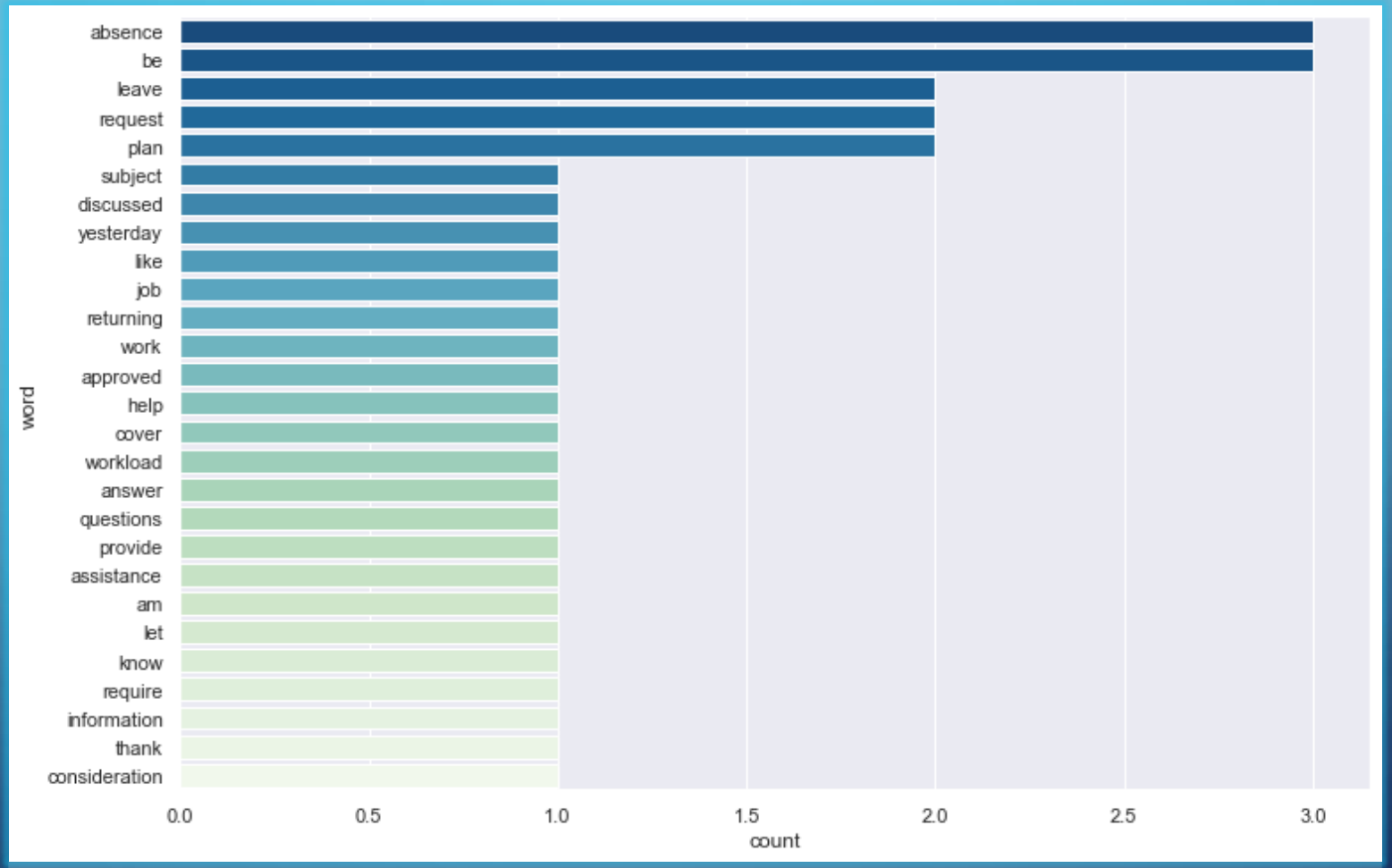
Tokenization is done to separate each and every word, space and punctuations used in the extracted text. Later the stop words are removed and a list containing the nouns and verbs are generated

```
*List without any Stopwords*:
 [
, Subject, :, Leave, Absence, -, John, Dooley,
, Dear, Jennifer, ,,
, discussed, yesterday, ,, like, request, formal, leave, absence, job, ., plan, away, August, 31, ,, 2020, October, 30, ,, 2020, ,, returning, work, November, 2, ,, 2020, .,
, approved, ,, glad, help, plan, cover, workload, absence, ., available, answer, questions, provide, assistance, away, .,
, let, know, require, additional, information, ., Thank, consideration, request, .,
, Best, ,,
, John]
```

Top Ten Word Count(Noun & Verb):

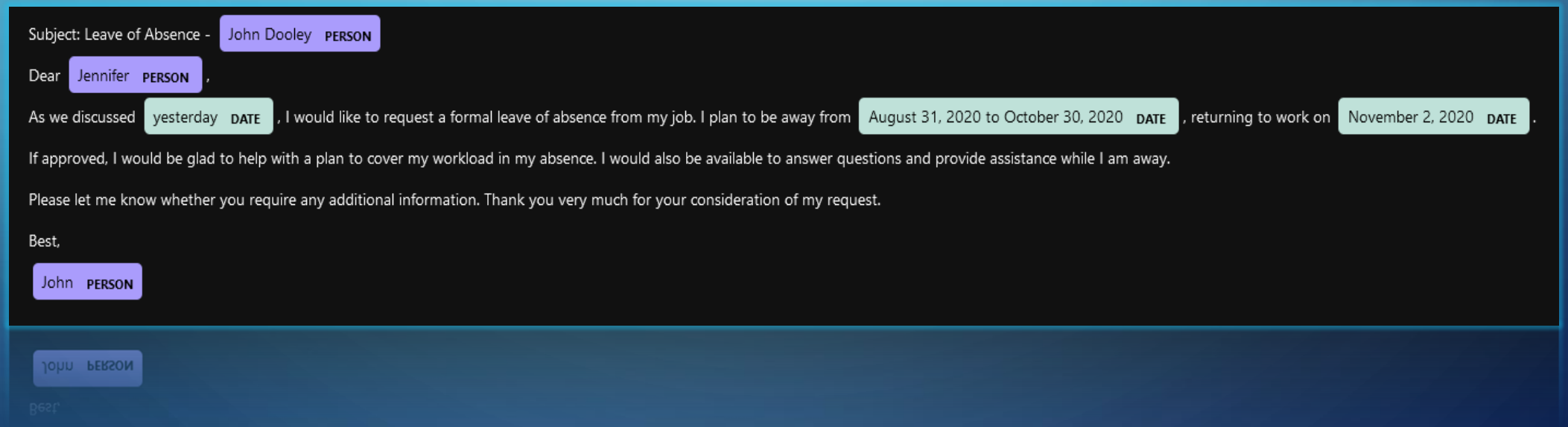
	word	count
0	absence	3
1	be	3
2	leave	2
3	request	2
4	plan	2
5	subject	1
6	discussed	1
7	yesterday	1
8	like	1
9	job	1

EDA (CONT.…)


absence
be
leave
request
plan
subject
discussed
yesterday
like
job
returning
work
approved
help
cover
workload
answer
questions
provide
assistance
am
let
know
require
information
thank
consideration
word
0.0
0.5
1.0
1.5
2.0
2.5
3.0
count

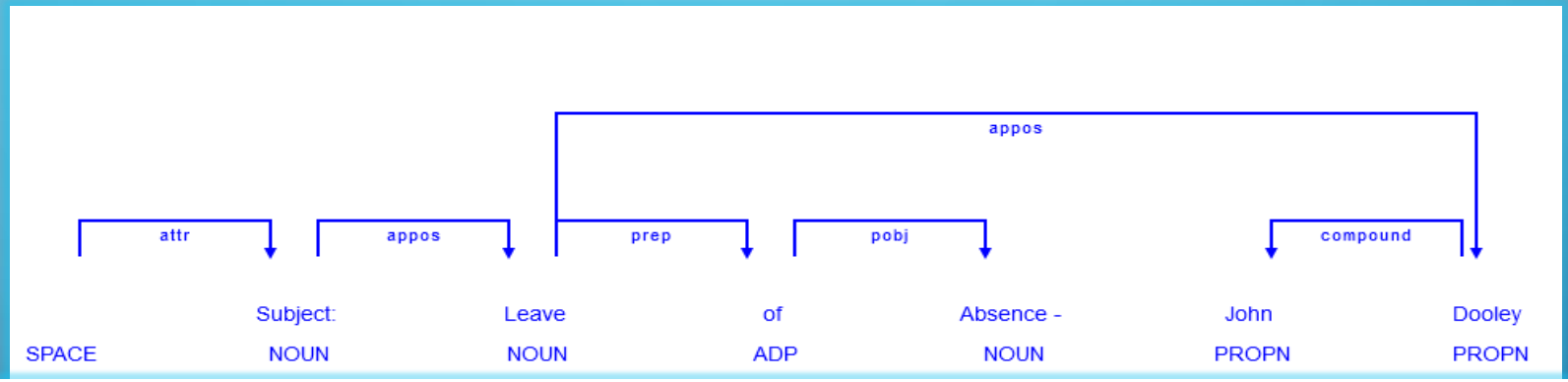
EDA (CONT.…)

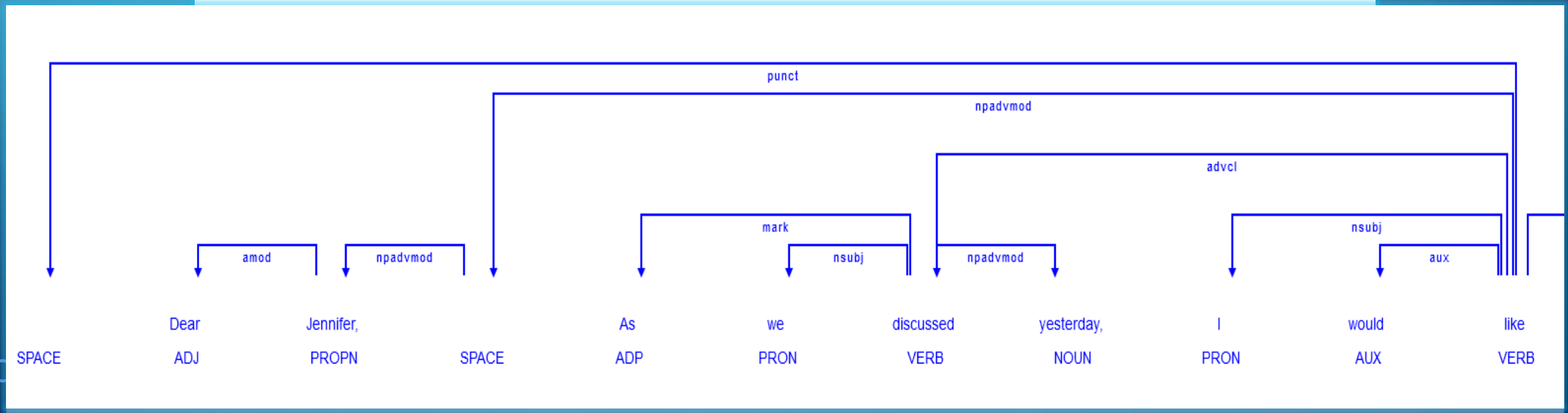
Named Entity Recognition(NER) is done on the extracted text and visualized as given below



EDA (CONT.…)

Dependency Tree or Parse Tree is visualized as given below. This establishes a link on how the tokens are related to each other





NLP MODEL

We have used an NLP library called **Spacy** and utilized the English language model, ***en_core_web_trf,*** as it had the highest accuracy evaluation metric when compared to the other English language models available

PSEUDOCODE

#Collect input from user

```
print(" 1. Vacation Leave Email Template
    2. Sick Leave Email Template
    3. Birthday Wishes Email Template
    4. Cover Letter Email Template
    5. Employee work appreciation Email Template
    6. Out of Office Email Template
    7. Thank you note for Business Email Template")
val = input("Enter your desired category(1-7) for the Email Template:>>> ")
```

#Based on the user selection, collect input for the respective keyword required

```
print("Keywords: absence, discussed, vacation, formally, approved, assistance, endeavored,
trip")
keyword = str(input("Please enter a keyword from the above list to generate the
Template:>>>"))
```

#Based on the keyword input, retrieve the respective content from the URL, using the div_id or classname

```
extract = Webscrape_divID(URL1, div_id1)
```

#Apply Spacy model on the extracted content

```
nlp = spacy.load('en_core_web_trf')
spacy_text = nlp(extract)
```

PSEUDOCODE (CONT….)

#EDA

```
Word_Frequency(spacy_text)
```
#Lists the count of words

```
Token_Attributes(spacy_text)
```
#Prints the token attributes such as index, isalpha, is punctuation, is stopword and shape of the token

```
POS_Attributes(spacy_text)
```
#Lists the attributes of POS tagged words

```
Macro_Visualize(spacy_text)
```
#Displays Parse tree/Dependency Tree(on the whole text)

```
POS_Tag(spacy_text)
```
#Displays POS Named Entities from the extracted text

#Generate Template from the extracted text

```
template = FindnReplace(spacy_text)
```
#Replaces the specific words with the generic ones and formats the text

```
print("****Here is your Template for the Email category you chose****", template)
```

GENERATED EMAIL TEMPLATE(SAMPLE)

****Here is your Template for Employee Work Appreciation Email****

Subject Line : Thank You Very Much !

Dear [Your Colleague's Name] ,

I wanted to let you know how much I appreciated your help with the [Work of Appreciation] .
I know how much time and effort you invested to not only get the [Work of Appreciation] done prior to the deadline , but to ensure the client was satisfied with every step of the process .
You are a valued member of our team , and I truly appreciate your contributions !
Best Regards ,

[Your Name]
[Your Designation]

DEPLOYMENT

- Our code was deployed as a web app using the Streamlit tool
- We provide a drop down menu to the end user to select their desired Email category
- Also we request the user to select a keyword set, to generate the Email template
- Based on the user selection, we display the various attributes and visualizations pertaining to the extracted text and finally the Email template is generated.

EMAIL TEMPLATE GENERATION

Using Webscraping and NLP Techniques to Generate an E-mail Template

Email Template Generation by Web-Scraping

Select your desrired category for the Email Template:

Drop down options

Vacation Leave Email Template

- Vacation Leave Email Template
- Sick Leave Email Template
- Birthday Wishes Email Template
- Cover Letter Email Template
- Employee work appreciation Email Template
- Out of Office Email Template
- Thank you note for Business Email Template

Select your desrired category for the Email Template:

Drop down options

Thank you note for Business Email Template

You have selected: Thank you note for Business Email Template

Please select the keyword set from the below options to generate the Template

Keyword-set Selection

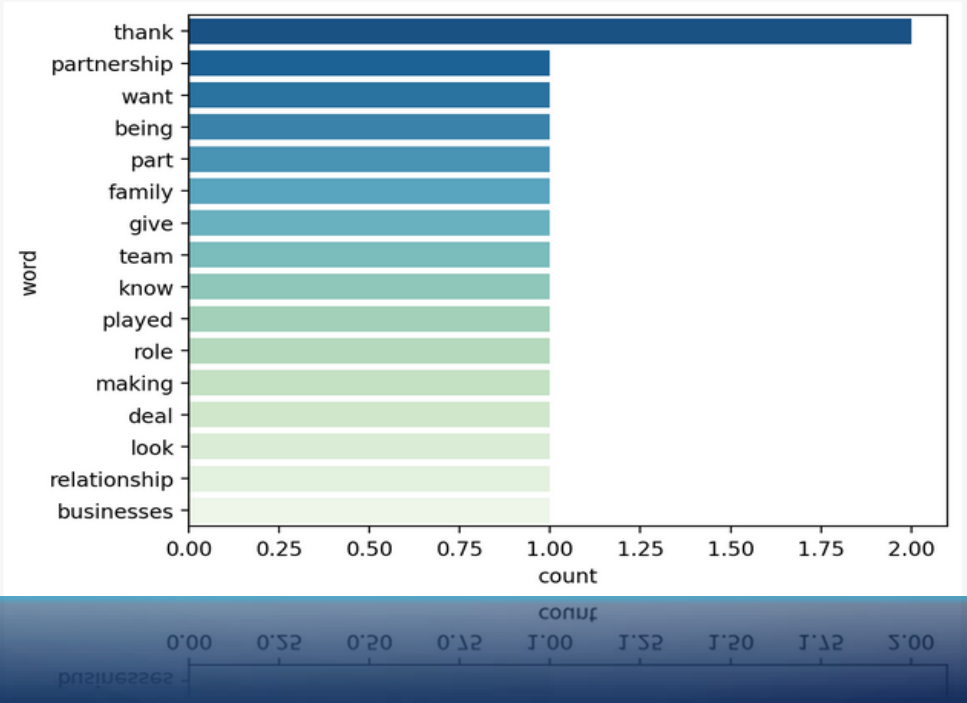
- (•) partnership, family, fruitful, business
- ( ) delighted, customer, loyal, patronage

You selected: partnership, family, fruitful, business

```
Scraping Web......Done!
```

Exracted text to be analyzed:

Dear Sam,I'm very excited about our new partnership, and I just want to thank you for being a part of our extended family. Please give my best to everyone on your team, as I know that they played no small role in making this deal possible. I look forward to a long and fruitful relationship between our two businesses. Thank you, again!Sincerely, Jonathan



Named Entities

Select entity labels +

Dear Sam PERSON , I'm very excited about our new partnership, and I just want to thank you for being a part of our extended family. Please give my best to everyone on your team, as I know that they played no small role in making this deal possible. I look forward to a long and fruitful relationship between our two CARDINAL businesses. Thank you, again!Sincerely, Jonathan PERSON

	text	label_	start	end	start_char	end_char
0	Sam	PERSON	2	3	6	9
1	two	CARDINAL	66	67	305	308
2	Jonathan	PERSON	74	75	349	357

Token attributes

Select token attributes +

	text	pos_	dep_	ent_type_
0		SPACE	ROOT	
1	Dear	ADJ	amod	
2	Sam	PROPN	npadvmod	PERSON
3	,	PUNCT	punct	
4		SPACE	ccomp	
5	I	PRON	nsubj	
6	'm	VERB	ROOT	
7	very	ADV	advmod	
8	excited	ADJ	acomp	
9	about	ADP	prep	
10	our	PRON	poss	

Your Template for Business Deal Closure Email

Dear [Your Partner's Name],

I 'm very excited about our new partnership and I just want to thank you for being a part of our extended family .
Please give my best to everyone on your team as I know that they played no small role in making this deal possible . I look forward to a long and fruitful relationship between our two businesses .
Thank you , again!
Sincerely,

[Your Name]
[Your Contact Number]

CHALLENGES

Below are a couple of challenges we came across

- There aren't enough Email Samples available on the web for different categories. We mainly get the Email Templates. As our project goal was to generate an Email Template from the Email Samples, we could not have considered the Templates directly from the web. Hence we have limited the template generation to two per category
- We cannot generalize our algorithm to automatically web scrape every URL and to generate the email template. As each and every webpage is designed differently, manual intervention is required to inspect the webpage and to fetch the data. Also, content formatting has to be done manually as each extracted content uses different formats


ngiyabonga
danke
謝謝
teşekkür ederim
спасибо
dank je
thank you
gracias
tapadh leat
mochchakkeram
bedankt
hvala
mauruuru
dziękuję
go raibh maith agat
obrigado
sagolun
sukriya
kop khun krap
grazie
arigatō
takk
dakujem
мерси
terima kasih
감사합니다
merci