

Capstone Project Weekly Progress Report

Semester	Fall 2021
Course Code	AML 3406
Section	Section 2
Project Title	Restaurant Recommending Chatbot
Group Name	Group 3
Student names/Student IDs	Vignesh Kumar Muruganathan C0793760
Reporting Week	Week 5 – 10/09/2021
Faculty Supervisor	Vahid Hadavi

1. Tasks Outlined in Previous Weekly Progress Report

- Created GIT Repository
- Downloaded the latest data from Yelp in Json format
- Loaded them in local repository
- Installed MongoDB in cloud service
- Imported the dataset into MongoDB
- Structured the data

2. Progress Made in Reporting Week

- Analyzed Business data and Review data
- Explored the data to find various insights in them
- Grouped data based on the restaurant category and visualized them in order to understand how the data is represented.

Data Collection

```
In [3]: business_df = pd.read_json('D:\Downloads\Yelp\yelp_academic_dataset_business.json',lines=True)
business_df.head()

Out[3]:
```

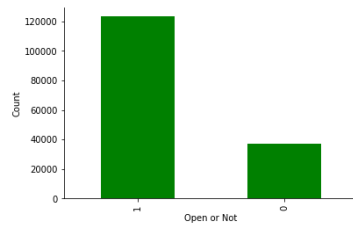
	business_id	name	address	city	state	postal_code	latitude	longitude	stars	review_count	is_open	attrib
0	6Yb2HFdyym3zuRgubtw	Oskar Blues Taproom	921 Pearl St	Boulder	CO	80302	40.017544	-105.283348	4.0	86	1	('RestaurantsTakeOut', 'True', 'WiFi'
1	1CddRFPZA0oIY5mHG3U0w	Flying Elephants at FDX	7000 NE Airport Way	Portland	OR	97218	45.588906	-122.593331	4.0	126	1	('RestaurantsTakeOut', 'T Restaurants'
2	bnN78MBNLprQ1a1y5dRq	The Reclamory	4720 Hawthorne Ave	Portland	OR	97214	45.511907	-122.613693	4.5	13	1	('BusinessAcceptsCreditCa True', 'Rest'
3	oaepeysvcUj17qwldcOVig	Great Clips	2566 Enterprise Rd	Orange City	FL	32763	28.914482	-81.295979	3.0	8	1	('RestaurantsPriceRange2 BusinessAcc'
4	PE8uqAjdwi0E4-8mjG3vVA	Crescent Terminus	1048 Memorial Dr SE	Atlanta	GA	30316	33.747027	-84.353424	4.0	14	1	('GoodForKids', 'Is BusinessParking'

Exploratory Data Analysis

```
: # To show number of open and closed businesses
open_bus = business_df.groupby('is_open').business_id.count()
open_bus

: is_open
0    37337
1    123248
Name: business_id, dtype: int64

: # To plot a bar graph to show the total number of open and closed businesses
business_df.is_open.value_counts(dropna=False).plot(kind='bar',color='green')
plt.xlabel('Open or Not')
plt.ylabel('Count')
plt.show()
sns.despine()
```



3. Difficulties Encountered in Reporting Week

- We faced few issues with the performance, so currently in choosing the best model which makes the model to perform better with more data.

4. Tasks to Be Completed in Next Week

- NLP Concepts - Finding the best restaurants for particular cuisines
- Modelling
- Model Evaluation
- Cosine Similarity
- Chat bot Building

Overall Project Plan:

Restaurant Recommending Chatbot

TASK NAME	RESPONSIBLE	START	FINISH	DURATION (DAYS)	STATUS	STATUS
Project Proposal						Complete
Case study Analysis	Team	18-Sep	23-Sep	5	In Progress	Overdue
Requirements Documentation	Vignesh/Murali	18-Sep	23-Sep	5	In Progress	In Progress
Presentation Slides	Swathi	18-Sep	23-Sep	5	In Progress	Not Started
Requirements Gathering						
S/W Environment Setup	Team	23-Sep	1-Oct	8	Not Started	
Data acquisition - Sorapy, Twitter API	Vignesh/Varadha	23-Sep	1-Oct	8	Not Started	
Requirement Analysis	Swathi	23-Sep	1-Oct	8	Not Started	
Documentation	Swathi/Murali	23-Sep	1-Oct	8	Not Started	
Development Phase I						
Model building	Vignesh/Murali	1-Oct	22-Oct	21	Not Started	
NLP	Varadha	1-Oct	22-Oct	21	Not Started	
Development Phase II						
Webapp Building	Swathi	23-Oct	12-Nov	20	Not Started	
Buffer	Team	23-Oct	12-Nov	20	Not Started	
Testing Phase I						
Unit Testing	Murali	13-Nov	19-Nov	6	Not Started	
Integration Testing	Swathi	20-Nov	26-Nov	6	Not Started	
Testing Phase II						
Functional Testing	varadha/murali	26-Nov	4-Dec	8	Not Started	
Performance Testing	vignesh/swathi	5-Dec	13-Dec	8	Not Started	
Final Presentation						
Report Generation/Documentation	Swathi/Murali	13-Dec	15-Dec	2	Not Started	
Slides data collection	Varadha/vignesh	15-Dec	16-Dec	1	Not Started	
Finalize Presentation	Team	16-Dec	17-Dec	1	Not Started	
Approve Presentation	Team	17-Dec	18-Dec	1	Not Started	

Individual Project plan:

Case study analysis

S/W environment setup

Development phase 1 – NLP

Development Phase 2 – Webapp Building

Testing Phase 2 – UNIT and Integration testing

HIGHLIGHTED BELOW ARE MY RESPONSIBILITIES

Restaurant Recommending Chatbot

TASK NAME	RESPONSIBLE	START	FINISH	DURATION (DAYS)	STATUS
Project Proposal					
Case study Analysis	Team	18-Sep	23-Sep	5	In Progress
Requirements Documentation	Vignesh/Murali	18-Sep	23-Sep	5	In Progress
Presentation Slides	Swathi	18-Sep	23-Sep	5	In Progress
Requirements Gathering					
S/W Environment Setup	Team	23-Sep	1-Oct	8	Not Started
Data acquisition - Scraping, Twitter API	Vignesh/Varadha	23-Sep	1-Oct	8	Not Started
Requirement Analysis	Swathi	23-Sep	1-Oct	8	Not Started
Documentation	Swathi/Murali	23-Sep	1-Oct	8	Not Started
Development Phase I					
Model building	Vignesh/Murali	1-Oct	22-Oct	21	Not Started
NLP	Varadha	1-Oct	22-Oct	21	Not Started
Development Phase II					
Webapp Building	Swathi	23-Oct	12-Nov	20	Not Started
Buffer	Team	23-Oct	12-Nov	20	Not Started
Testing Phase I					
Unit Testing	Murali	13-Nov	19-Nov	6	Not Started
Integration Testing	Swathi	20-Nov	26-Nov	6	Not Started
Testing Phase II					
Functional Testing	varadha/murali	26-Nov	4-Dec	8	Not Started
Performance Testing	vignesh/swathi	5-Dec	13-Dec	8	Not Started
Final Presentation					
Report Generation/Documentation	Swathi/Murali	13-Dec	15-Dec	2	Not Started
Slides data collection	Varadha/vignesh	15-Dec	16-Dec	1	Not Started
Finalize Presentation	Team	16-Dec	17-Dec	1	Not Started
Approve Presentation	Team	17-Dec	18-Dec	1	Not Started