**CSE 6363 001 Machine Learning**

**Fall 2022**

**Project 2 - Source Code Instructions**

**Instructor's Name: Dr. Dajiang Zhu**

**Student Name: Siva Srinivasa Sameer Miriyala**
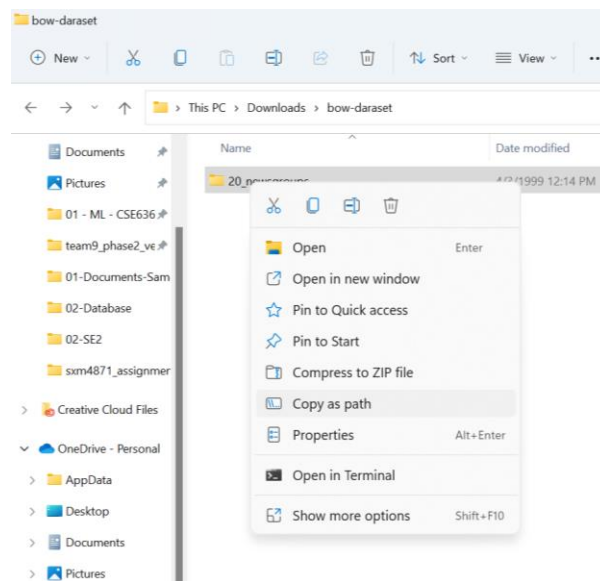
**Student ID: 1002024871**

**Instructions to run the program:**

**Step-1:** Open the .ipynb file in yout Jupyter Notebook.

**Step-2:** Kindly update the path in the source code in your local system where the dataset is actually present. In my case the dataset was present in the following path:

```python
data_path = 'C:/Users/samee/Downloads/bow-daraset/20_newsgroups'
data_dict, news_groups_dict = load_data(data_path)
stopwords_list = load_corpus()
wf_dict = {}
```

**Step-3:** How to copy the path ? Right click on the **'20_newgroups'** folder and click on copy path

After you copy the path, Please make sure to change the backward slashes to the forward slashes.

For Example:

The below path **will not work** . Please see the next step.
*Before*: **"C:\Users\samee\Downloads\bow-daraset\20_newsgroups"**

The below path will work properly. Replace the below path in the code
*After*: **"C:/Users/samee/Downloads/bow-daraset/20_newsgroups"**

**Step-4:** After replacing the path in the code. Run the .ipynb file. This will take around 15mins to execute as the loading of the data will take some time.