**CSCE-5222 Feature Engineering**

In class task 1 (ICE-1)

Please goto <https://towardsdatascience.com/text-classification-in-python-dd95d264c802> and follow the article.

This article is for text classification using python.

Their Github is available at <https://github.com/miguelfzafra/Latest-News-Classifier/tree/master/0.%20Latest%20News%20Classifier>.

Follow their step 00,01,02,03 and 04.

After successfully implementing their code. Try to gather data from an online URL related to autonomous cars (your choice but a long article)

Use all techniques covered in the above code on the dataset that you have just created.

**Submission:**

The submission of the task would be done to GitHub and Canvas. GitHub should have source code, images, and WIKI. WIKI page of GitHub should have all code embedded and photos should be inline, also if there are any questions asked in the exercise they should be included in the wiki. Canvas submission should have source code file and your wiki contents copied in the word file.

**Policy:**

In class task can be performed in teams of 3 -5. However, the submission is individual to your respective GitHub and Canvas. The source code can remain the same but the answers to the questions should be different from each other as the answers will testify your concept regarding the topic and you will be graded according to that.

**Grading:**

|  |  |  |
| --- | --- | --- |
| No. | Item Description | Percentage |
| 1 | Complete and proper Github Submission | 10% |
| 2 | Complete and proper submission to Canvas | 5% |
| 3 | Source Code | 50% |
| 4 | Explaining the answers | 30% |
| 5 | Commenting, formatting, and visualizing your code properly and timely submission | 5% |
|  | Total | 100% |