**Project Instructions file**

* In the first step , I have gone through the objectives as per the assignment , then I decided to click and go on those links to check weather they are working or not in my PC properly because, sometimes in my PC whenever I go for some new links my PC firewall will refuse to load links but , luckily it was successfully done .
* In the second step , I have decided to to extract the data from the given links one by one , but then after some time I realized that it was time taking process . Then I decided to extract the data using loops , for which I have loaded the dataset into the PC . To load all the dataset I have to import the Padas library , for data extraction , I need to import the BeautifulSoup library form bs4 .
* Then , imported those libraries and start extracting data from the given links present in the Input.xlsx dataset.
* I used for loops and putting all the articles with their titles into the new separate files by giving their names on their url\_ids.
* For which , now I have imputed all the text files which are required for data cleaning given into the zip files , then I used loops to remove all the stop words from the articles
* Then , I have started importing all the positive and negative text files given the zip file named as Master dictionary.
* Then , I started checking all the positive and negative words present in the articles on the basis of them all the other variables are filles like Polarity , Subjectivity , Average Word count , fog index for checking readability of the article .
* In the same way I approached all the variables.
* Then , I have put them into the dataset one by one column wise for each article and , then imported the Output data structure file and start imputing all the variables which are null in the Output data structure file.
* Run .py file in the vs code , to generate the output , ,ipynb file is more preferable in the case where we have do analysis , but no worry I have given comments that what is going all the codes , and you will easily understand it.
* PC should contain powerful Operating System , like Windows , LINUX .
* Your PC should be of i3 , i5 or o7 generation and contains .
* You have to use VS CODE to run the code You want the file In the .py , but , if you are using .ipynb then , use jupyter notebook to run the code in the block of cells.
* You have to import the pandas , numpy , nltk.tokenize , BeautifulSoup from bs4 libraries in to your virtual environment or , also import it into your command prompt as well to use it in the entire PC .
* Use command pip install pandas , pip install numpy , pip install nltk , pip install bs4 libraries are required to be used.
* IDELLIC should be installed of the latest version download ir easily from the google.