INSY 5378: Homework 3 (Extra) – Text Analyzer

**DUE DATE**: Oct 26th, Wednesday by 11:59 p.m.

This is the third (extra) Python assignment. The total is 100 points. As this is an extra assignment, non-submission won’t hurt your final grade.

Instructions:

1. Late submissions will not be accepted.
2. Submit your Python code (no output file please) in Blackboard. DO NOT email your homework

This assignment is an extension from HW 2, where we analyzed text files. Instead of using the imperative programming approach, we want to re-implement them in **a class**.

Your goal is to develop a **class** Textfile that provides methods to analyze a text file.

1. [20 points] The class Textfile will support a constructor that takes as input a file name (as a string) and instantiate a Textfile object associated with the corresponding text file. You may want to create instance variables to keep the text, words, lines, and unique words for later.
2. [20 points] The Textfile class should support methods nchars(), nwords(),and nlines() that **return** the numbers of characters, words, and lines, respectively, in the associated text file.
3. [20 points] The class should also support methods read() and readlines() that **return** the content of the text file as a string or as a list of lines, respectively, just as we would expect for file objects.
4. [20 points] The class should support method grep() that takes a target string as input and searches for lines in the text file that contain the target string. The method **returns** the list of lines in the file containing the target string; in addition, the method should **print** the line number and the line. The line numbering starts at 0.
5. [20 points] Finally, the class should support method words() that takes no input and **returns** a list, without duplicates, of words in the file.

You would have to create a file name “textfile.py” and work from there. You would need to implement the following methods:

1. \_\_init\_\_()
2. nchars(), nwords(), nlines()
3. read(), readlines()
4. grep()
5. words()

Below you can find a sample run using ‘raven.txt’ and ‘frankenstein.txt’ files.

>>> from textfile import \*

>>> t = Textfile(‘raven.txt’)

>>> t.nchars()

6299

>>> t.nwords()

1125

>>> t.nlines()

126

>>> print(t.read())

Once upon a midnight dreary, while I pondered weak and weary,

…

Shall be lifted nevermore!

>>> t.grep(‘nevermore’)

75: Of “Never-nvermore.”

89: She shall press, ah, nevermore!

124: Shall be lifed – nevermore!

['Of "Never-nevermore."\'', 'She shall press, ah, nevermore!', 'Shall be lifted - nevermore!']

>>>

>>> s = Textfile(‘frankenstein.txt’)

>>> s.nchars()

421098

>>> s.nwords()

74999

>>> s.nlines()

7441

Demo video is here:

<https://www.youtube.com/watch?v=2WJ7YJX51gg>

**Note:** Just submit your Python code (textfile.py) in Blackboard. You don’t need to submit any output files.