

Assignment 3 Marking Guide

Robin Dawes

October 15, 2021

Correctness : 85%

1. Loads the set of Stop Words from a file. 10%
2. Lets the user choose the directory where the text documents are located. 10%
3. Accesses all the files in that directory with ".txt" extension. 10%
4. For each document, determines the set of words that represents it. 20%
5. For each document, applies the Jaccard Similarity measure to determine which of the other documents has the greatest similarity to the one being considered 20%
6. Presents the results in tabular form using tkinter's grid layout. 15%

Programming Style : 15%

1. Appropriate naming : 5%
2. Appropriate organization of code : 5%
This includes:
 - function definitions located as specified in Style Guide
 - good use of functions to isolate minor and/or repetitive tasks
 - good use of data structures such as dictionaries, sets, tuples
3. HTML documentation of program and functions generated by pydoc : 5%