

Information Storage and Retrieval Course, Year 2022/2023

Task (3)

- 1- Add a fourth choice to task (2) and when the user chooses choice number 4 ask the user again to choose which stemmer he/she prefers: porter stemmer or snowball stemmer.

Choice number 1: print tokenized words

Choice number 2: print tokenized sentences

Choice number 3: print original text.

Choice number 4: stemming

Choice 4.1: porter Stemmer

Choice 4.2: Snowball Stemmer

Example (1):

Input: programming

Input: Choice number 4 & porter stemmer

Output: program

Example (2):

Input: programming

Input: Choice number 4 & snowball stemmer

Output: programming

Bonus:

Remove null rows from dataset and print the number of null rows.

Remarks:

- Due Date is Sunday 19/3/2023.
- Solve this task **individually**.