

MLDS 422 – Fall 2023
Homework 1
Due Friday, 9/29/23 at 11:59pm

Python Practice

Your task is to build a word counter using base Python only (no libraries like numpy, pandas, collections, etc.).

More details are in the Jupyter Notebook called **Homework_1.ipynb**.

1. Write a function called **count_words_basic** that takes in the test phrase and returns a dictionary of word counts.
2. Write a function called **count_words** that takes in the test phrase and returns a dictionary of word counts.
3. Write a function called **most_common** that returns the most common word from a string. Apply the function to the test phrase. In your solution, explain what the code is actually doing.
4. Write a function called **count_words_from_file** that allows you to read in a file, clean the text, count the words and return the most common word. Apply the function to the `peter_pan_chapter_1.txt` file.
5. Expand on the previous functions to create a function called **count_words_in_many_documents** that reads in a list of documents, and creates a dictionary of word counts across all the documents. Apply the function to the list of documents in the Jupyter Notebook.
6. Create a function called **return_top_n_words** that takes in a dictionary of word counts (output of previous question) and returns the top n words specified by the user. The default value of n should be 3 if the user doesn't specify in a value. Apply the function to the output of the previous question.