

Workshop 11

Applications POS and NER

Task 1

Implement a program `workshop11-task1.py` to

1. Read the text file (The Adventures of Sherlock Holmes).
 2. Determine the top 20 noun chunks that are most mentioned in this story along with their frequencies in descending order.
 3. For each named entity of types person, organization and location, count the number of occurrences of each name and display the top 10 occurrences of each entity type in descending order.
- Submit your code on MS Teams by 1:30pm.

Task 2

Write a program `worksheet1 1-task2.py` to analyze the given Harry Potter books.

1. Devise a way to determine the leading characters in each book. Implement your algorithm accordingly.
2. Implement an algorithm to list the 15 important places in each Harry Potter book.
 - Analyze your algorithm. Are you happy with the results? Write your analysis and submit it in PDF along with your code.
 - Submit your work on MS Teams by 1:30pm.