

кафедра комп'ютерних технологій, ДНУ
2022/2023

Тема: «Інтелектуальний аналіз тексту за допомогою WordCloud»

Постановка задачі:

Розглядається метод Machine Learning – для інтелектуального аналізу тексту за допомогою WordCloud. Для оцінки тональності (емоційного забарвлення) тексту. Цей метод використовується для швидкого та наочного аналізу великих текстів та листів.

Створити звіт та описати процес вирішення наступної задачі WordCloud:

- 1) Встановити інтерпретатор мови Python та IDE
- 2) Створити текстовий файл з даними - (текст на ваш розсуд, у прикладі Far Over the Misty Mountains Cold – з книжки The Hobbit)
- 3) Використовуючи мову Python зчитати файл та **ВИКЛЮЧИТИ** з нього:

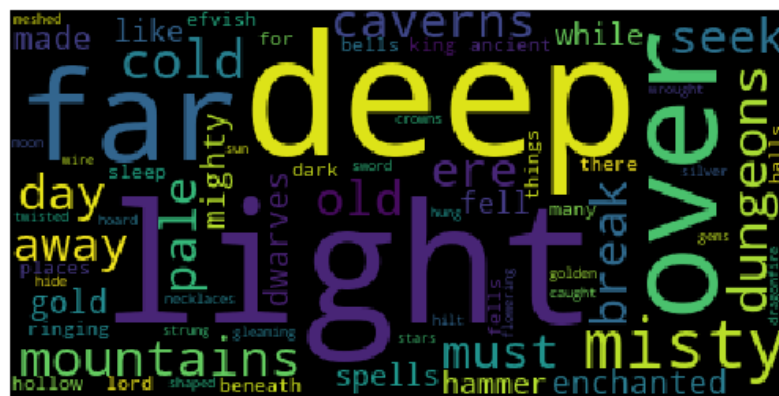
Знаки пунктуації:

""!()-[]{};:'"\<>./?@#\$%^&* ~""

Нецікаві для аналізу слова:

["the", "a", "to", "if", "is", "it", "of", "and", "or", "an", "as", "i", "me", "my",
"we", "our", "ours", "you", "your", "yours", "he", "she", "him", "his", "her", "hers",
"its", "they", "them", "their", "what", "which", "who", "whom", "this", "that", "am",
"are", "was", "were", "be", "been", "being", "have", "has", "had", "do", "does", "did",
"but", "at", "by", "with", "from", "here", "when", "where", "how", "all", "any", "both",
"each", "few", "more", "some", "such", "no", "nor", "too", "very", "can", "will", "just",
"on", "in"]

- 4) Створити словник (асоціативний масив в інших мовах - dict, map) зі слів та їх входжень:
Dict<String Word, Int Count>
- 5) Відобразити ХмаруСлів – WordCloud



Додатково:

- Можна використовувати різні бібліотеки. У даному прикладі використовуються:
`from matplotlib import pyplot as plt`
`from os import path`
`import wordcloud`
- Можна використовувати і інші мови програмування, але найпростіше вирішити цю задачу саме на мові Python.