Міністерство освіти і науки України

Західноукраїнський національний університет

Факультет комп’ютерних інформаційних технологій

Кафедра інформаційно-обчислювальних систем і управління

Звіт до лабораторної роботи №11

З дисципліни “Інтелектуальний аналіз даних”

Студента групи КН-31

Ковальковського В. В.

Тернопіль 2023

**Лабораторна робота №11**

**Визначення емоційного забарвлення новин**

Варіант:

| 13 | 14 | ABP Majha | Marathi |

**Парсинг**

Збір даних (назв статей) я вирішив проводити по тегу <p> та класі «other-news», так як в цих параметрах прописувались назви статей



Рис. 1 – Результат виконання порсингу

**Переклад назв**

Переклад я вирішив зробити в два кроки, спочатку з маратхі на англійську, потім з англійської на українську



Рис. 2 – Перекладені назви на англійську

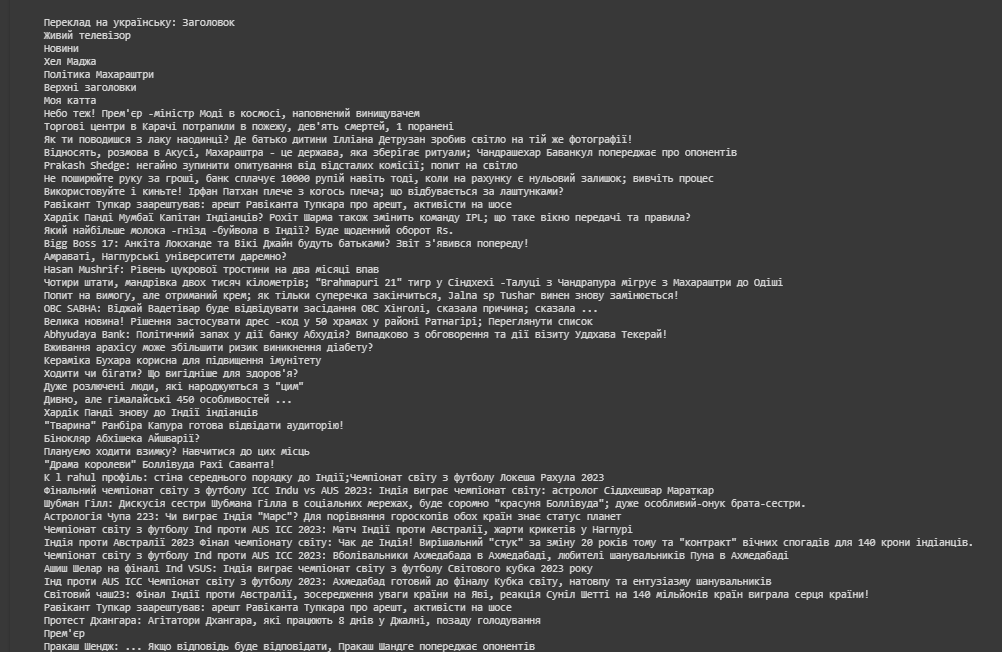


Рис. 3 – Перекладені назви на українську

**Простий аналіз слів**

Проводимо видалення символів, які не є кирилецею, подвійних пробілів. Також токенізуємо текст і видаляємо стоп-слова.



Рис. 4 – Токенізований текст

Виділяємо унікальні слова



Рис.5 – Унікальні слова

Розтавляємо емоційну оцінку для цих унікальних слів

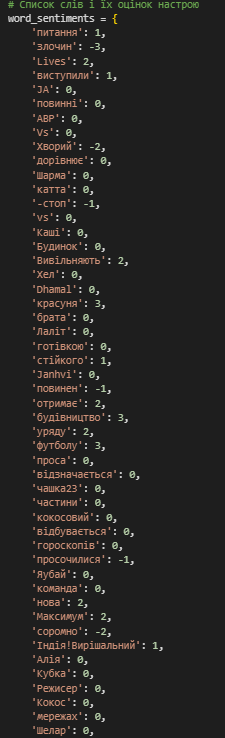


Рис. 6 – Емоційна оцінка унікальних слів

Результат емоційної оцінки:

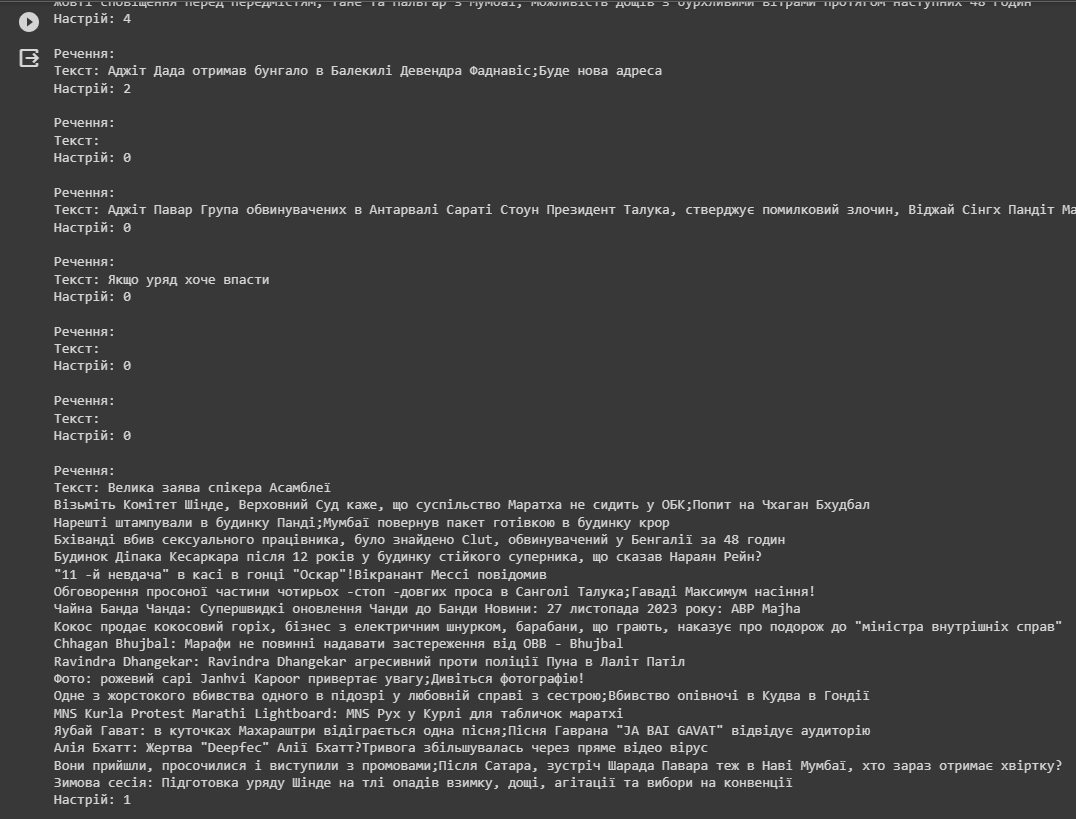


Рис. 7 – Весь текст

Результат емоційної оцінки, де настрій набуває якогось значення:

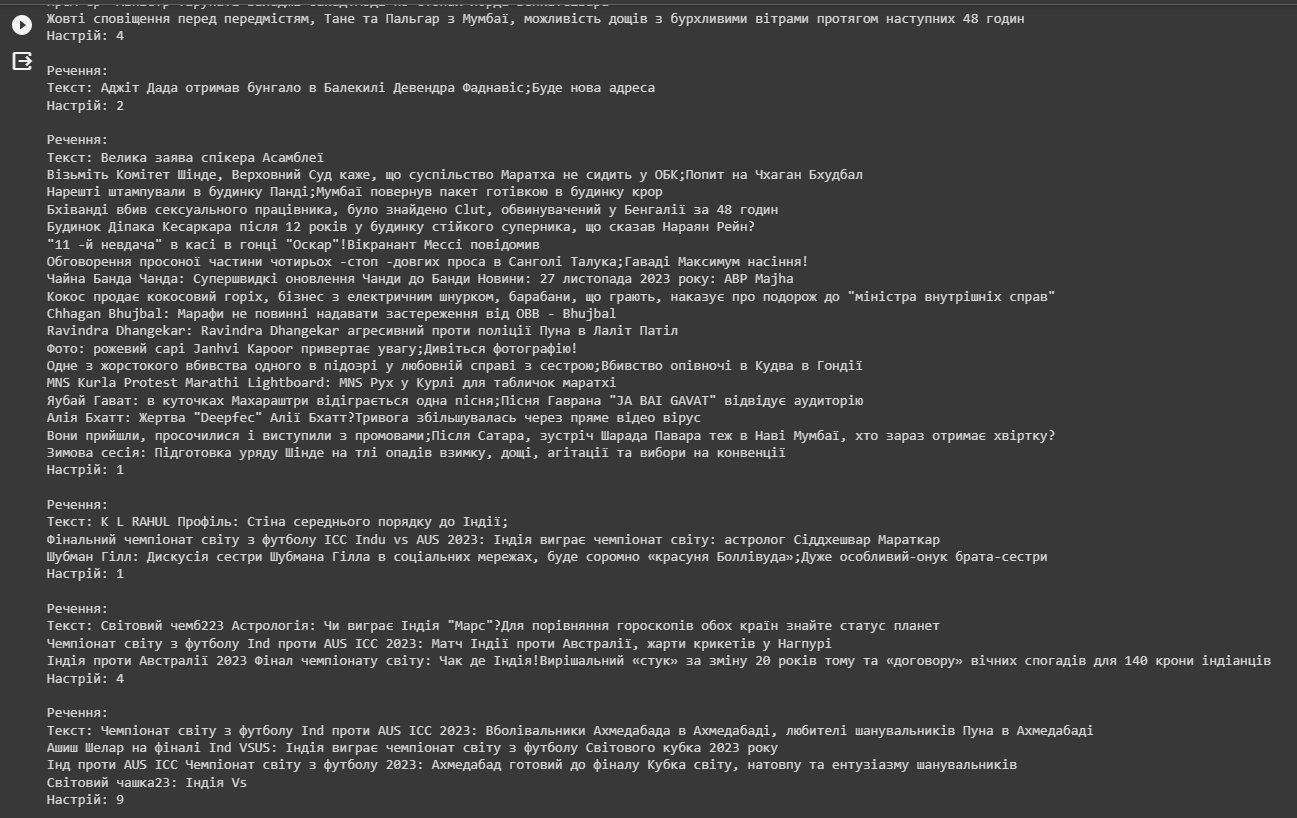


Рис. 8 – Назви статей, де емоційна оцінка набуває відмінного від 0 значення

Висновок: під час виконання лабораторної роботи, був проведений парсинг сайту новин, простий аналіз отриманих з парсингу даних та аналізування отриманих результатів за допомогою емоційної оцінки