МІНІСТЕРСТВО ОСВІТИ ТА НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ОДЕСЬКА ПОЛІТЕХНІКА»

Кафедра інформаційних технологій

Звіт

Модульна контрольна робота №1

З дисципліни: «МАОВД»

Варіант 1

Виконав:

здобувач групи АД-222

Буберенко С.О.

Перевірив:

Рудніченко М.М.

Оцінка\_\_\_\_\_\_\_\_\_\_\_\_

Одеса 2024

Питання

1. Поняття про дані і їх призначення, філософський підхід
2. Які існують базові принципи обробки «Великих даних»?
3. З якою метою і яким чином використовується шар компонентів системи «Подання результату»?
4. З яких компонентів складається обладнання для обробки «Великих даних»?

Відповіді

1. Поняття про дані і їх призначення, філософський підхід

Дані — це інформація, що фіксується для подальшого аналізу та прийняття рішень. Філософськи дані розглядаються як "сировина" знання, основа для побудови розуміння світу, що потребує інтерпретації. Їх призначення — забезпечення істинності тверджень через досвід і аналітичні підходи.

2. Базові принципи обробки «Великих даних»

- Обсяг (Volume): обробка великих масивів даних.

- Швидкість (Velocity): висока швидкість обробки та передачі.

- Різноманітність (Variety): обробка даних різних типів (структуровані, неструктуровані).

- Достовірність (Veracity): забезпечення точності та надійності даних.

3. Шар «Подання результату»

Цей шар використовується для візуалізації та представлення оброблених даних користувачеві у зрозумілому вигляді. Він включає звіти, графіки, дашборди і дозволяє сприймати результати аналізу для прийняття рішень.

4. Компоненти обладнання для обробки «Великих даних»

- Сервери з високою обчислювальною потужністю.

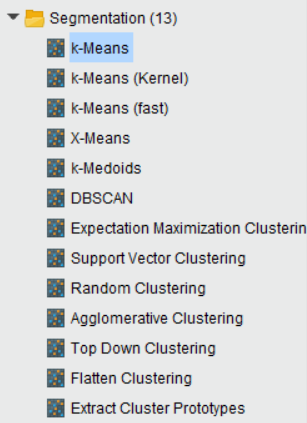
- Системи зберігання даних (HDD, SSD, хмарні сховища).

- Мережеві інтерфейси для швидкої передачі даних.

- Програмне забезпечення для аналізу та управління даними (Hadoop, Spark).

Практична частина

Використовуючи приклад кластеризації на базі системи Rapidminer реалізувати рішення задачі кластеризації документів засобами іншої системи **(за вибором Orange, Knime чи на базі написання коду у Jupiter Notebook)** (самі вхідні документи сформувати самостійно вручну в txt файлах із статей/журналів/книг або сайтів, дані один у одного не копіювати) на базі використання 3 різних алгоритмів, взятих зі списку:



Приклад уважно дивимося, наведено нижче.

Внаслідок реалізації надати до протоколу:

1. Архів із файлами даних

2. Файли реалізації

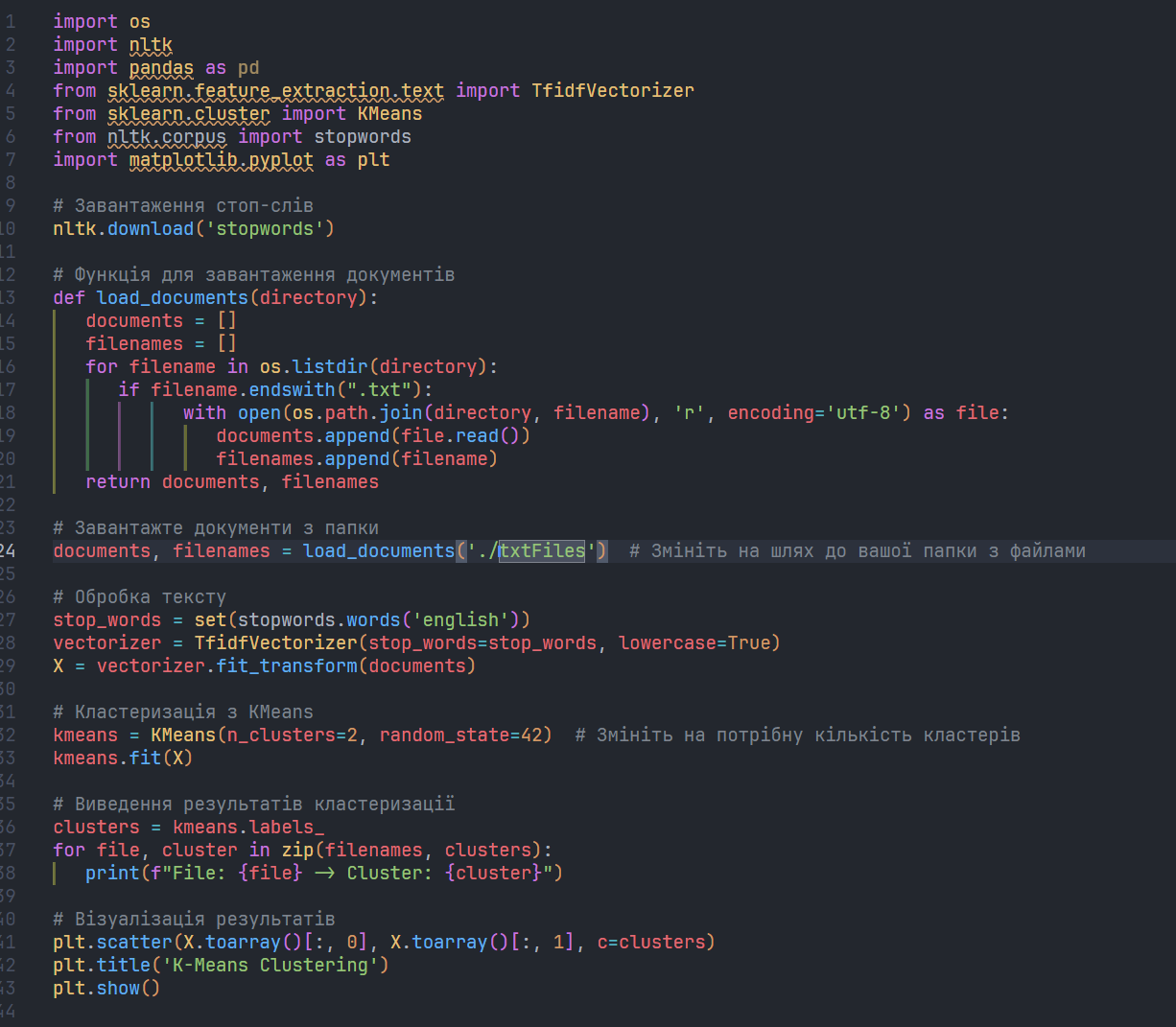


Рисунок 1. Код программи