

Аналіз отриманих результатів.

Втілений проект полягав у тому, щоб створити швидкий пошук українською мовою та використовувати для зберігання даних структури даних. З аналізу, зробленого у попередньому домашньому завданні, вимоги відповідають моїм очікуванням.

Результати проекту не лише показують швидкість знаходження відео, а також важливість нормалізації запиту та субтитрів, використання структур даних для зберігання потрібних даних про відео.

Фактор, який впливає на отримані результати, це спосіб зберігання даних. Оскільки у мене є до 2000 файлів із субтитрами, то я змогла використати структури даних для «індексування» та обрахування довечності відео за формулами, описаними у зареферованих дописах. Проте при такому підході із збільшенням кількості відео, збільшується час пошуку, адже при кожному запуску програми рахується доречність кожного відео відносно конкретного запиту. Якщо обсяг даних набагато більший, то потрібно або записувати частоти слів у файл, або використовувати бази даних.

Говорячи про структури даних, варто зазначити тип, який я використовувала для свого проекту. Хоча вибір структур даних був доволі широким, я зупинилася на одно- та двозв'язних списках. Таким чином у мене усі необхідні дані для виведення на веб-сторінку зберігаються в одному місці та їх генерація проводиться саме під час пошуку. Я могла також обрати масив, проте у цій реалізації проекту я користувалася списком, який є набагато зручніший під час нормалізації даних, адже мені доводилося працювати переважно зі стрічками.

Ще одним важливим фактором, який сприяє доречності знайдених відео, це нормалізація даних. Хоча і без неї пошук все одно знайде результати, вони не будуть найкращими. Видаляючи слова без лексикографічного значення, а також проведення стемінгу збільшує ймовірність того, що результати будуть доречними.

Також я проводила аналіз доречності знайдених відео. Зайшовши на сторінку [ted.com](https://www.ted.com) я вводила перекладений запит англійською мовою та порівнювала результати. Хоча на ці тематики могли або бути додані відео, яких немає у csv файлі, який я використовувала, або ж вони не містили українських субтитрів, знайдені відео моєю програмою знаходилися в перших п'яти відео, що відповідає вимогам доречності знайдених відео.

Проте є частина мого проекту, яка мені вдалася не найкраще. Хоча мені доводилося створювати веб-сторінку, цього разу мені вперше прийшлося створювати її дизайн. Аудит моєї веб-сторінки показав середній результат через те, що вона не завжди добре адаптується під різні розміри дисплеїв.

Висновки до циклу домашніх завдань

Цикл домашніх завдань був розрахований на використання структур даних в обраному проекті, а також переконатися, що вони є оптимальним рішенням для зберігання даних. Метою цього циклу домашніх завдань було навчитися розробляти власний проект, навіть і початковий, адже я змогла не лише написати розв'язок до завдання, а ще його сформулювати та спроектувати план роботи. Це виявився цікавий та корисний досвід для мене.

У даному проекті втілені усі вимоги з вищим пріоритетом: на сторінку виводиться опис, кількість переглядів, назва і т.ін. знайденого відео, а також youtube посилання на б інших відео. Також є можливість переглядати відео локально. Протягом розроблення програми було зауважено, що пошук відбувається швидше, якщо запускати програму локально, а не з `pythonanywhere` сервера. Спричинено це, перш за все, різною швидкістю Інтернету. Тому якщо запускати програму локально, то пошук здійснюється в середньому за 2-3 секунди, що є в два рази менше, ніж час, зазначений у нефункціональних вимогах. Проект, як і планувалося, розміщений на pythonanywhere.com та використано Flask фреймворк для створення веб-сторінки.

Проте були й вимоги, які я не втілила у своєму проекті. Завантаження субтитрів знайдених відео не вдалося втілити через те, що субтитри знаходяться на моєму акаунті на `pythonanywhere`, які я отримала через `арі`, та виникає помилка при їх завантаженні. Також одними з вимог було розміщення всіх можливих файлів з субтитрами на веб-сторінку. Ця вимога не могла бути виконана через причину, зазначену вище.

Отже, незважаючи на те, що я не змогла виконати усіх функціональних вимог до свого проекту, усі нефункціональні вимоги та функціональні вимоги з вищим пріоритетом мені втілити вдалося. Загалом я задоволена результатом роботи, адже пошук відбувається швидко, а результат пошуку переважно доречний (винятком може слугувати те, що на таку тему взагалі немає відео). Також важливою деталлю є те, що пошук відбувався саме по субтитрах, а не по назві, що робить його значно ґрунтовнішим.