

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
“КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
ІМЕНІ ІГОРЯ СІКОРСЬКОГО”**

Факультет прикладної математики
Кафедра програмного забезпечення комп'ютерних систем

**КУРСОВИЙ ПРОЕКТ
ТЕХНІЧНЕ ЗАВДАННЯ**
з дисципліни “Бази даних”

спеціальність 121 – Програмна інженерія

на тему: _____
(назва теми)

Студент
групи КП-81

Дикий І. М.

(підпис)

Викладач
к.т.н, доцент кафедри
СПіСКС

Петрашенко А.В.

(підпис)

Київ – 2021

Найменування та галузь застосування розробки

Найменування: моніторингова система ресурсів оренди житла в курортних зонах Українських Карпат: фільтрація, зберігання та аналіз вартості житла залежно від його стану/виду і розміщення з метою оцінки певних тенденцій.

Галузь застосувань: спрощення розрахунку приблизної вартості мандрівки для туристів та моніторингу ситуації ринку для виставлення власної ціни для власників нерухомості.

Дата початку та закінчення проекту

Дата початку проекту - 05.03.2021.

Дата закінчення проекту - 20.05.21.

Мета розробки

Метою розробки даної моніторингової системи є набуття виконавцем практичних навичок розробки сучасного програмного забезпечення, що взаємодіє з постреляційними базами даних, а також здобуття навичок оформлення відповідного текстового, програмного та ілюстративного матеріалу у формі проектної документації.

Вимоги до програмного забезпечення

1. Генерація даних

Буде реалізовано стягування даних про житло у курортних зонах карпат з декількох основних ресурсів, а саме: booking.karpaty.ua, booking.com, doba.ua.

2. Фільтрації та валідації даних

Буде реалізовано перевірку даних на дублювання шляхом порівняння назв та локацій закладів та на перевірку на коректність.

3. Реплікація та масштабування

Будуть використовуватися у MongoDB для реплікації Replica Sets а для масштабування шардинг.

4. Аналіз даних

Буде використовуватися бібліотека JSAT(Java Statistical Analysis Tool). В рамках аналізу буде проводитися структурування інформації з різних ресурсів, валідація, знаходження взаємозв'язку між вартістю, місцем та сезоном,

5. Резервування та відновлення даних

Це включає у себе вищезазначених Replica Sets.

Обґрунтування вибору СУБД

В даній моніторинговій системі буде використовуватись MongoDB. Це документоорієнтована СУБД, не потребує жорсткої моделі даних, що пришвидшить і спростить розробку зараз та в подальшому. Окрім цього дана СУБД дозволяє зберігати дані в об'єктному вигляді, що в більшій мірі і відповідає нашій моделі даних. Відповідно це пришвидшить додавання та отримання даних. Також ця СУБД підтримує горизонтальне масштабування з метою зменшення навантаження на кожен окремий вузол шляхом розподілення навантаження між ними всіма.

Вимоги до інтерфейсу користувача

Інтерфейс даної системи буде консольним. Він буде використовуватися для:

- отримання параметрів пошуку та аналізу
- запуск пошуку та аналізу
- відображення звітної інформації
- генерації графіків у вигляді збережених на диск файлів
- перегляду конкретних пропозицій

Вибір засобів розробки

Для реалізації даної системи буде використовуватися мова програмування Python 3.8 оскільки це є проста мова програмування пристосована до подібного роду задач, яка в свою чергу є надійною і крос-платформеною. До того ж наявна велика кількість бібліотек для збору і аналізу даних. Зокрема будуть використовуватися наступні технології:

- ***numpy*** - математична бібліотека мови Python, що додає підтримку великих багатовимірних масивів і матриць, разом з великою бібліотекою високорівневих математичних функцій для операцій з цими масивами.
- ***pandas*** - це бібліотека на Python для обробки і аналізу даних. Робота pandas з даними будується поверх бібліотеки NumPy.
- ***matplotlib*** - це бібліотека Python 2D, яка представляє числові дані у різноманітних форматах та інтерактивних середовищах на різних платформах
- ***folium*** - це бібліотека мови Python для візуалізації даних на інтерактивній мапі.
- ***scrappy*** - це швидкий веб-фреймворк з відкритим вихідним кодом, написаний на Python, який використовується для отримання даних з веб-сторінки за допомогою селекторів на основі XPath.

- *pymongo* - є дистрибутивом Python, що містить інструменти для роботи з MongoDB, і є рекомендованим способом роботи з MongoDB від Python.

Етапи розробки

№	Етап розробки системи	Термін виконання
1	Затвердження теми курсової роботи. Розробка технічного завдання	5.03.2021
2	Розробка засобів генерації даних	16.03.2021
3	Розробка засобів фільтрації та валідації даних.	24.03.2021
4	Реалізація зберігання, реплікації та масштабування інформації розроблюваної системи.	01.04.2021
5	Додавання засобів аналізу даних	10.04.2021
6	Додавання засобів резервування та відновлення даних	18.04.2021
7	Розробка користувацького інтерфейсу	23.04.2021

8	Тестування системи	28.04.2021
9	Виправлення помилок	04.05.2021
10	Підготовка матеріалів курсового проекту та оформлення пояснювальної записки	12.05.2021
11	Захист курсової роботи	20.05.2021