**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: **Моніторингова система онлайн магазину**

**Студент**

**групи** КП-73 Труш Артем Олександрович

(підпис)

**Викладач**

**к.т.н, доцент кафедри**

**СПіСКС Петрашенко А.В.**

(підпис)

Київ – 2020

**ЗМІСТ**

стор.

[1. НАЙМЕНУВАННЯ ТА ГАЛУЗЬ ЗАСТОСУВАННЯ РОЗРОБКИ 3](#_Toc34787594)

[2. ДАТА ПОЧАТКУ ТА ЗАКІНЧЕННЯ ПРОЕКТУ 3](#_Toc34787595)

[3. МЕТА РОЗРОБКИ 3](#_Toc34787596)

[4. ВИМОГИ ДО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ 3](#_Toc34787597)

[5. ОБҐРУНТУВАННЯ ВИБОРУ СУБД 5](#_Toc34787598)

[6. ВИМОГИ ДО ІНТЕРФЕЙСУ КОРИСТУВАЧА 5](#_Toc34787599)

[7. ВИБІР ЗАСОБІВ РОЗРОБКИ 5](#_Toc34787600)

[8. ЕТАПИ РОЗРОБКИ 6](#_Toc34787601)

1. НАЙМЕНУВАННЯ ТА ГАЛУЗЬ ЗАСТОСУВАННЯ РОЗРОБКИ

Найменування: моніторингова система онлайн магазину.

Галузь застосування: планування розташування точок видачі чи магазинів в залежності від перебування покупців, що дозволили використовувати їх геолокацію; на основі замовлень робити прогнози щодо визначення який товар потрібен покупцю, який найпоширеніший, який час найсприятливіший для покупки якого товару, в залежності від змін ціни; робити висновки, чи є в продукту конкуренти, чи варто створити аналог, що зможе конкурувати з цим продуктом.

2. ДАТА ПОЧАТКУ ТА ЗАКІНЧЕННЯ ПРОЕКТУ

Дата початку проекту: 26.02.2020

Дата закінчення проекту: 21.05.2020

3. МЕТА РОЗРОБКИ

Метою розробки є створення програмного забезпечення, яке буде взаємодіяти з нереляційною базою даних, проводити дослідження і подальший аналіз, оцінку, порівняння отриманих результатів для виявлення певних закономірностей, тенденцій, змінних і їх динаміки. А також здобуття навичок використання NoSQL баз даних та інструментальними засобами підтримки розробки додатків для подібних баз даних, оформлення відповідного текстового, програмного та ілюстративного матеріалу у формі проектної документації.

4. ВИМОГИ ДО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

* Вимоги до генерації даних:

Дані будуть формуватися рандомізатором, так як немає можливості отримати актуальні дані про використання. Будемо вважати, що ці дані – дані певного магазину техніки, подібного Цитрусу. Має бути сформовано не менш ніж 50000 користувачів, більшість з яких поділилась своїм розташуванням, які використовували сайт протягом певного часу і просто розглядали певну техніку чи робили замовлення. Точки координат будуть генеруватись в межах Києва, на основі них виділяти райони, в яких необхідно розмістити магазин. Замовлення будуть формуватися також випадковим чином.

* Вимоги до фільтрації та валідації даних:

Щоразу при отриманні нових даних активність клієнта слід фільтрувати наступним чином: якщо користувач зайшов на сайт, та перебував там менше 3 хвилин, не враховувати дані про місцеположення цього користувача; людина рухається, в такому випадку, якщо буде певна область, в яку втрапило найбільше точок координат обрати будь-яку точку з цієї області, якщо не буде, то першу точку; Також потрібно фільтрувати користувачів, які користуються сайтом з других міст.

* Вимоги до засобів реплікації та масштабування:

оскільки для використання у даній розробці була обрана нереляційна база даних MongoDB, то як засіб реплікації використовуватимемо реплісети (Replica Sets), а для масштабування шардинг (sharding), який використовує MongoDb для обробки великої кількості даних.

* Вимоги до аналізу даних:

Отримані дані повинні бути проаналізовані створеним програмним забезпеченням в наступний спосіб: в залежності від дня тижня та часу провести аналіз, в який період часу найбільше покупців здійснюють покупки, щоб можна було створювати акції, в дні коли найменша кількість покупців. Провести оцінку кількості часу, що покупець витрачає на вибір товару і оцінити зручність інтерфейсу в залежності від його активності. Проаналізувати, яку техніку часто купують, на якій техніці найбільше схвальних відгуків, щоб розуміти скільки і якої техніки потрібно замовити. В залежності від отриманих точок геолокації виділяти райони, де необхідно розмістити магазини і в яких треба це зробити в першу чергу(спираючись на те, скільки техніки куплено і на яку суму купують). Для реалізації алгоритмів використовувати математичні бібліотеки або методи Big Data тощо.

* Вимоги до резервування та відновлення даних:

передбачені при використанні Replica Sets у MongoDb.

5. ОБҐРУНТУВАННЯ ВИБОРУ СУБД

Була обрана найпопулярніша серед нереляційних СКБД - MongoDb. Це документо-орієнтована система керування базами даних із відкритим кодом, яка не потребує опису схеми таблиць. MongoDB займає нішу між швидкими і масштабованими системами, що оперують даними у форматі ключ/значення, і реляційними СКБД, функціональними і зручними у формуванні запитів.

Вибір нереляційної СКБД обгрунтовується наявністю великої кількості ненормалізованих даних, які необхідно обробити швидко. Вона забезпечує можливість отримання неприведених до норм даних та подальшу роботу із ними. Формування додаткових таблиць при появі додаткової інформації у екземплярі (реляційні СКБД) було б недоцільним.

6. ВИМОГИ ДО ІНТЕРФЕЙСУ КОРИСТУВАЧА

Інтерфейс користувача виконати в консольному варіанті. Повинні бути доступні наступні розділи (пункти в меню) в інтерфейсі: налаштування засобів та підсистем (генерації даних, масштабування, резервування та відновлення даних), запуск/завершення їх роботи, генерація звітної інформації (графіків, діаграм) у вигляді збережених файлів-зображень. Звітна інформація стосується візуалізації роботи засобів аналізу даних.

7. ВИБІР ЗАСОБІВ РОЗРОБКИ

У якості засобів розробки було обрано мову Python 3.6. СКБД**:** MongoDB. Для аналізу даних повинні використовуватися відповідні бібліотеки напрямку Data Science чи Big Data, такі, як: numpy, pandas, matplotlib, scipy.

8. ЕТАПИ РОЗРОБКИ

|  |  |  |
| --- | --- | --- |
| **№** | **Назва етапів розроблення** | **Термін виконання** |
| 1 | Затвердження теми курсової роботи. Опрацювання відповідної літератури. Розроблення та узгодження технічного завдання. | 12.03.2020 |
| 2 | Аналіз постановки задачі | 22.03.2020 |
| 3 | Розробка засобів генерації даних. | 05.04.2020 |
| 4 | Додавання засобів фільтрації та валідації даних. | 12.04.2020 |
| 5 | Реалізація зберігання, реплікації та масштабування інформації розробленої моніторингової системи. | 19.04.2020 |
| 6 | Додавання засобів аналізу даних. | 03.05.2020 |
| 7 | Додавання засобів резервування та відновлення даних | 10.05.2020 |
| 8 | Тестування програми | 15.05.2020 |
| 9 | Аналіз результатів. Підготовка матеріалів курсового проекту та оформлення пояснювальної записки | 17.05.2020 |
| 10 | Захист курсової роботи | 21.05.2020 |