

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

Факультет прикладної математики
Кафедра програмного забезпечення комп'ютерних систем

МЕТОДИЧНІ ВКАЗІВКИ
до виконання курсового проекту з дисципліни

“Бази даних”

для студентів напряму підготовки 121 – Програмна інженерія

Укладач: к.т.н., доцент А.В.Петрашенко

Київ - 2020

ЗМІСТ

ЗМІСТ	2
Загальні положення	3
Загальні вимоги	3
Тематика курсового проекту	3
Загальна структура програмних засобів	5
Склад та оформлення курсового проекту	7
Вимоги до технічного завдання	7
Вимоги до програмного забезпечення	8
Вимоги до бази даних	8
Обґрунтування вибору СУБД	8
Вимоги до інтерфейсу користувача	8
Вибір засобів розробки	8
Етапи розробки	9
Вимоги до пояснювальної записки	9
Титульний аркуш	9
Анотація	10
Зміст	10
Вступ	10
Аналіз інструментарію для виконання курсового проекту	10
Аналіз функціонування засобів масштабування	10
Опис результатів аналізу предметної галузі	10
Висновки	10
Література	11
Додатки	11
Важливі дати	11
Критерії оцінювання	12
ДОДАТОК 1	13
ДОДАТОК 2	14

Загальні положення

Курсовий проект (КП) з дисципліни “Бази даних” передбачає виконання комплексного завдання щодо створення програмного забезпечення, орієнтованого на роботу із базою даних постріляційного типу NoSQL (XML, «ключ-значення», документо-орієнтовані, графові тощо).

Метою КП є набуття студентами практичних навичок розробки сучасного програмного забезпечення, що взаємодіє з постріляційними базами даних, а також здобуття навичок оформлення відповідного текстового, програмного та ілюстративного матеріалу у формі проектної документації.

У результаті виконання курсового проекту студенти повинні вміти розробляти програмне забезпечення для постріляційних баз даних, володіти основами використання СУБД, а також інструментальними засобами підтримки розробки додатків для подібних баз даних.

Загальні вимоги

Робота виконується студентом самостійно після затвердження викладачем технічного завдання. Забезпечити можливість демонстрації роботи програмних засобів на двох комп'ютерах.

Тематика курсового проекту

Тематика курсового проекту визначається у співпраці викладача зі студентом на основі запропонованого останнім технічного завдання.

Об'єктом розробки у курсовому проекті є моніторингова система широкого призначення. Під *моніторингом* слід розуміти проведення низки однотипних замірів досліджуваного об'єкта і подальший аналіз, оцінку, порівняння отриманих результатів для виявлення певних закономірностей, тенденцій, змінних і їх динаміки.

Моніторингову систему пропонується розробити для наступних предметних галузей:

1. Моніторингова система медичного призначення: можливе замірювання показників роботи органів людини з метою прийняття рішення у випадку виходу показників за межі норми.
2. Моніторингова система кліматичних показників (довкілля, приміщення тощо): безперервне замірювання показників температури (води, повітря), вологості повітря, рівня шкідливих речовин у повітрі (екологічний моніторинг).
3. Моніторингова система в галузі організації дорожнього руху: безперервне зчитування зображень доріг або номерних знаків автомобілів з метою подальшого зберігання та аналізу (заторів, пошуку автомобілів з певним номерним знаком тощо).
4. Моніторингова система новинних повідомлень в Інтернет: фільтрація, зберігання та аналіз новинних потоків у Інтернет з метою оцінки певних трендів та тематично пов'язаних ланцюгів новин.
5. Моніторингова система за цінами споживчих товарів: фільтрація, зберігання та аналіз поточних цін на товари за певною категорією з метою визначення певних цінових трендів, передбачення популярності товарів тощо.
6. Моніторингова система за рухом громадського транспорту: безперервний контроль наявності на маршруті машин громадського транспорту з метою оцінки відповідності розкладу руху, можливого виходу з ладу або аварій на маршруті.
7. Моніторингова система спостереження за критичним об'єктом: безперервний контроль за допомогою сенсорних пристроїв, розміщених на визначеній території і призначених для відслідковування зміни стану деяких об'єктів: руху, звукових сигналів, хімічних викидів тощо.
8. Моніторингова система контролю успішності учнів (студентів) навчального закладу: збір, фільтрація та аналіз оцінок учнів різних вікових категорій за різними напрямками підготовки з метою прогнозування та вдосконалення навчального процесу.
9. Моніторингова система мережного трафіку комп'ютерної мережі: збір, фільтрування пакетів мережного трафіку з метою визначення

підозрілих з точки зору інформаційної безпеки або виходу з ладу певних вузлів мережі.

10. Моніторингова система Web-сайтів: збір, фільтрація та відслідковування відвідування сторінок та розділів Web-сайту (наприклад, соціальної мережі) з метою визначення шляхів удосконалення представлення інформації, створення додаткових розділів тощо.

Примітка. Студенту дозволяється запропонувати власну предметну галузь, що розглядається, та перелік задач, що планується розв'язати (аналогічну запропонованим вище).

Загальна структура програмних засобів

Програмні засоби мають містити наступні компоненти (див. рис.1).

1. Підсистема попередньої обробки даних, що складається з:

- 1.1. Засоби генерації даних. Мають забезпечувати генерацію адекватних предметній галузі псевдовипадкових даних у потоковому та пакетному режимах. Рекомендується для реалізації даної задачі розробити окрему утиліту, яку можна буде запустити у вигляді декількох екземплярів для емуляції наявності багатьох джерел даних, що одночасно функціонують. Як варіант генерації даних можна використати відкриті ресурси Інтернет, наприклад, за допомогою бібліотеки Scrapy.

- 1.2. Засоби фільтрації та валідації даних. Мають забезпечувати можливість знаходження “корисних” для даної задачі даних та перевірку їх на відповідність заданим структурам даних та вимогам до них.

2. База даних системи призначена для зберігання, реплікації та масштабування інформації розробленої моніторингової системи.

3. Засоби реплікації та масштабування входять у склад сервера бази даних і призначені для забезпечення цілісності, узгодженості та доступності даних,

що зберігаються у СУБД. Вибір СУБД (або декількох) залишається за студентом.

4. Засоби аналізу даних можуть включати реалізацію алгоритмів і методів машинного навчання, Data Mining, Big Data тощо. Метою цих засобів є виявлення прихованих залежностей між даними для розв'язання задач прогнозування, визначення поточного стану, трендів тощо. **Увага!** Реалізацію алгоритмів необхідно запозичити у відомих бібліотеках та фреймворках, зокрема, OpenCV, numpy, scipy, pandas та інших.

5. Засоби резервування та відновлення даних призначені для оперативного та пакетного збереження фрагментів та всієї бази даних з можливістю її відновлення з урахуванням необхідності підключення додаткового комп'ютера як елемента горизонтального масштабування.

Примітка. За необхідності студент може розширити перелік функціональних можливостей, зокрема, засобами атрибутивного та повнотекстового пошуку тощо.



Рис.1. Узагальнена структура програмних засобів

Склад та оформлення курсового проекту

Курсовий проект складається з таких частин:

1. титульний аркуш (див. Додаток 1);
2. технічне завдання;
3. пояснювальна записка;
4. графічна частина;
5. програмне забезпечення.

Графічна частина виконується у вигляді креслень (плакатів) за допомогою графічних пакетів на аркушах формату А4.

Пояснювальна записка виконується на аркушах формату А4. Усі текстові і графічні документи зшиваються у одну книгу.

До захисту подаються книга текстових і графічних документів, а також електронні версії: програмного забезпечення, база даних, а також текстові та графічні документи на компакт-диску. Вихідні коди також мають бути опубліковані в репозиторії на сайті <https://github.com>.

Вимоги до технічного завдання

Технічне завдання є окремим документом, метою якого є визначення технічних та часових складових щодо створення програмного забезпечення курсового проекту.

У загальному вигляді технічне завдання повинно мати наступні розділи:

- титульний аркуш (див. додаток 2);
- найменування та галузь застосування розробки;
- дата початку та закінчення проекту;
- мета розробки;
- вимоги до програмного забезпечення;
- обґрунтування вибору СУБД;
- вимоги до інтерфейсу користувача;

- вибір засобів розробки;
- етапи розробки.

Вимоги до програмного забезпечення

У даному розділі технічного завдання необхідно конкретно до обраної предметної галузі описати що саме буде зроблено для реалізації кожного із структурних елементів системи, а саме засобів: генерації даних (зокрема, обґрунтування умов генерування псевдовипадкових значень), фільтрації та валідації даних, засобів реплікації та масштабування, аналізу даних (які задачі та можливі алгоритми будуть використані), резервування та відновлення даних.

Вимоги до бази даних

Генерування даних має відбуватись відповідно до вимог масштабування: а саме, має бути проілюстровано, що на двох комп'ютерах ефективність обробки та аналізу даних підвищується. Підтвердити ці положення результатами дослідження: таблицями, графіками, діаграмами.

Обґрунтування вибору СУБД

Дати коротку характеристику обраної СУБД, відзначити їх переваги для обраної задачі у порівнянні з аналогами (до 10 рядків).

Вимоги до інтерфейсу користувача

Інтерфейс користувача має бути мінімалістичним (достатнім буде реалізувати консольний варіант). Задачею інтерфейсу користувача є налаштування засобів та підсистем, запуск/завершення їх роботи, генерація звітної інформації (графіків, діаграм тощо) у вигляді збережених файлів-зображень. Звітна інформація стосується візуалізації роботи засобів аналізу даних (визначення трендів, вузьких місць тощо).

Вибір засобів розробки

У якості засобів розробки рекомендується використовувати мову Python 3.6 та відповідні бібліотеки напрямку Data Science. За необхідності студент може

запропонувати власний набір програмного інструментарію, обґрунтувавши його у технічному завданні.

Етапи розробки

У даному розділі необхідно у табличному вигляді подати детальний перелік етапів розробки, включно із розробкою технічного завдання, підготовкою кожного із обов'язкових структурних елементів програмного забезпечення (див. п. "Загальна структура програмних засобів"). Також необхідно вказати дату виконання кожного з етапів.

Вимоги до пояснювальної записки

Структурно пояснювальна записка складається з таких розділів:

1. Анотація.
2. Зміст.
3. Вступ.
4. Аналіз інструментарію для виконання курсового проекту.
5. Структура бази даних.
6. Опис програмного забезпечення.
 - 6.1 Загальна структура програмного забезпечення.
 - 6.2 Опис модулів програмного забезпечення.
 - 6.3 Опис основних алгоритмів роботи.
7. Аналіз функціонування засобів масштабування.
8. Опис результатів аналізу предметної галузі.
9. Висновки.
10. Література.
11. Додатки
 - А. Графічні матеріали
 - Б. Фрагменти програмного коду

Титульний аркуш

Титульний аркуш має бути встановленого зразку. Форма титульного аркуша наведена у додатку.

Титульний аркуш не нумерується як розділ, не вноситься до змісту і не нумерується як сторінка.

Анотація

В анотації у реферативному стилі наводиться інформація про зміст та результати, що отримані в курсовому проекті. Як розділ анотація не нумерується.

Зміст

Зміст повинен вміщувати в собі назви всіх розділів, підрозділів, пунктів та підпунктів, які є у курсовому проекті, а також перелік додатків. Визначення сторінок обов'язкове. Зміст характеризує структуру курсового проекту. Як розділ зміст не нумерується.

Вступ

У вступі на одній-двох сторінках описується призначення проекту, його актуальність, загальна постановка завдання. Вступ як розділ не нумерується.

Аналіз інструментарію для виконання курсового проекту

Аналіз літературних та джерел Інтернет виконується з метою визначення найбільш ефективного інструментарію щодо виконання курсового проекту і має включати порівняльний аналіз СУБД, бібліотек і фреймворків, що відповідають меті роботи.

Аналіз функціонування засобів масштабування

В даному розділі слід вказати, яким чином проводилось тестування масштабування: які обсяги інформації виявилось ефективним зберігати та аналізувати на одному комп'ютері, при яких граничних обсягах - на двох.

Опис результатів аналізу предметної галузі

В даному розділі слід проаналізувати результати виконання алгоритмів, що були використані в роботі. Дати текстовий опис, а результати у графічній формі надати у додатку.

Висновки

У висновках по проекту в реферативній формі повинні бути описані результати, отримані студентом на кожному з етапів виконання проекту, а

також висновки щодо досягнення мети курсового проекту тощо. Висновки пишуться на 1-2 сторінки і як розділ не нумеруються.

Література

Необхідно навести перелік джерел, на які були посилання в тексті. Список повинен формуватися в порядку посилань за текстом і вміщувати бібліографічні відомості офіційно виданих книжок, статей тощо.

Як розділ перелік літератури не нумерується. Формат опису літературних джерел повинен відповідати ДСТУ.

Додатки

Графічні матеріали, що наводяться у додатку мають включати:

- структуру бази даних за варіантом;
- Графіки, таблиці, діаграми як результат аналізу даних та дослідження масштабування системи;

Важливі дати

Підготовка технічного завдання на затвердження -- 12 березня 2020 року.

Захист курсового проекту -- 21 травня 2020 року.

Критерії оцінювання

Критерій	Максимальна кількість балів
Своєчасність подання ТЗ	5 (штраф: -1 кожен тиждень)
Засоби попередньої обробки (генерування та фільтрація даних)	10
Масштабування (шардінг) даних (включно із демонстрацією на двох вузлах)	15
Реплікація даних	10
Засоби статистичного аналізу даних	15
Засоби резервування/відновлення даних	10
Використання наукових бібліотек	15
Оформлення роботи	10
Своєчасність захисту роботи	5 (штраф: -1 кожен тиждень)
Засоби візуалізації результатів роботи (графіки, діаграми тощо)	5
Всього:	100

**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

з дисципліни “Бази даних”

спеціальність 121 – Програмна інженерія

на тему: _____
(назва теми)

Студент

групи КП-71(72,73)

(ПІБ)

(підпис)

Викладач

к.т.н, доцент кафедри

СПіСКС

Петрашенко А.В.

(підпис)

Захищено з оцінкою _____

Київ – 2020

**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-71(72,73)

(ПІБ)

(підпис)

Викладач

к.т.н, доцент кафедри

СПіСКС

Петрашенко А.В.

(підпис)

Київ – 2020