



Силабус курсу

Інструментальні засоби обробки даних

Ступінь вищої освіти – бакалавр

Галузь знань 12 “Інформаційні технології”

Спеціальність 123 «Комп’ютерна інженерія»

Освітньо-професійна програма: «Комп’ютерна інженерія»

Рік навчання: 3, Семестр: 6

Кредитів: 5 Мова викладання: українська та вірменська

Керівник курсу

ППП

Овакімян Оганес Арутюнович

Контактна інформація

erfiltanx@yahoo.com

Опис дисципліни

Мета дисципліни «Інструментальні засоби обробки даних» полягає в підготовці кваліфікованих фахівців, здатних ефективно використовувати сучасні інструменти та технології для збору, обробки, аналізу та візуалізації даних різних типів і розмірів. Студенти отримають компетенції, необхідні для роботи з великими масивами даних, їх оптимізації та трансформації з метою підтримки прийняття обґрунтованих рішень.

Завдання навчальної дисципліни спрямоване на вивчення фундаментальні концепції та принципів, які лежать в основі обробки і аналізу даних, з оглядом на сучасні методи та підходи, розвиток практичних навичок в області використання сучасних інструментів мов програмування платформ і технологій для ефективного обробки даних, вивчення технологій роботи з великими даними та робота з сучасними методами та засобами аналітики даних.

Структура курсу

Тема	Результати навчання	Завдання
Тема 1. Вступ в аналіз і обробку даних.	Розуміти поняття даних і концепцію роботи над ними. Проектування структури.	Питання, лабораторна робота
Тема 2. Запити за даними, фільтрування.	Розуміти та використовувати запити направлені на отримання даних і фільтрування даних.	Питання, лабораторна робота
Тема 3. Об’єднання таблиць: теорія та практика. Особливості типів даних у системах баз даних.	Знати типи об’єднань таблиць та задачі які вони вирішують. Володіти базовими типами даних.	Питання, лабораторна робота
Тема 4. Маніпуляція даними, групування елементів.	Знати техніки маніпуляції даними, групування елементів та сортування.	Питання, лабораторна робота
Тема 5. Транзакції та підзапити. Ефективність дотримання цілісності бази даних.	Знати типи транзакцій, розуміти як вони виконуються і які проблеми вирішують, розуміти де потрібно використовувати кожен з типів.	Питання, лабораторна робота

Тема 6. Обмеження для забезпечення цілісності та консистентності збережених даних.	Знати які обмеження використовуються для забезпечення цілісності і консистентності даних.	Питання, лабораторна робота
Тема 7. Бази даних та сховища даних.	Розуміти різницю між базою даних та сховищем даних.	Питання, лабораторна робота
Тема 8. Розподілені платформи.	Знати розподілені платформи та сфери застосування.	Питання, лабораторна робота
Тема 9. Інструментальні засоби.	Знати інструментальні засоби які використовуються для обробки даних.	Питання, лабораторна робота
Тема 10. Платформи обробки даних.	Знати платформи які надають можливості роботи з великими даними.	Питання, лабораторна робота
Тема 11. Методи запуску циклічних задач по обробці даних Apache Airflow.	Розуміти концепцію Airflow та циклічних задач, виконання таких задач у власному середовищі.	Питання, лабораторна робота
Тема 12. Методи перетворення даних у процесі ETL та ELT, DBT.	Розуміти концепцію DBT та процесу перетворення. Виконання таких процесів у власному середовищі.	Питання, лабораторна робота
Тема 13. Засоби візуалізації чистих даних: PowerBI, Looker.	Розуміти концепцію PowerBI та процесу візуалізації. Виконання таких процесів у власному середовищі.	Питання, лабораторна робота
Тема 14. PowerBI запити, підключення до зовнішніх джерел даних, перетворення та особливості агрегування даних.	Розуміти концепцію зв'язків у PowerBI та для чого вони використовуються. Розуміти роботу з зовнішніми джерелами даних.	Питання, лабораторна робота

У процесі вивчення даної дисципліни використовуються наступні способи оцінювання та методи демонстрування результатів навчання: поточне опитування, тестування; презентації результатів виконаних завдань; оцінювання результатів самостійної роботи студентів; інші види індивідуальних і групових завдань; залік.

Політика щодо академічної доброчесності: списування під час здачі заліків та екзаменів заборонено (в т.ч. із використанням мобільних девайсів).

Літературні джерела

Основна література

1. Data Engineering with dbt: A practical guide to building a cloud-based, pragmatic, and dependable data platform with SQL 1st Edition, Kindle Edition, Roberto Zagni, 2023.
2. Data Pipelines with Apache Airflow, Bas P. Harenslak, 2021.
3. Булига О. Згуртоване сімейство програм MS Power BI. Національний транспортний університет, 2022. 148 с.
4. Олещенко Л.М. Технології оброблення великих даних: конспект лекцій. Київ: КПІ ім. Ігоря Сікорського, 2021. 227 с

5. Підручник з Python. – URL: <https://docs.python.org/uk/3.13/tutorial/index.html>
6. Python for Data Analysis: Data Wrangling with pandas, NumPy, and Jupyter, 3rd Edition, Wes McKinney, 2022.
7. Zgurovsky M.Z., Zaychenko Y.P. Big Data: Conceptual Analysis and Applications. Springer, 2020. 298 p.

Додаткова література

8. Balamurugan Balusamy, Nandhini Abirami, Seifedine Kadry , Amir H. Gandomi . Big Data: Concepts, Technology, and Architecture. Wiley; 1st edition (2021). 368 p.
9. Neylson Crepalde . Big Data on Kubernetes: A practical guide to building efficient and scalable data solutions. Packt Publishing (2024). 296 p.
10. Гордійчук-Бублівська О. В. Методи та засоби опрацювання великих даних в розподілених інформаційних системах. Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 122 “Комп’ютерні науки”. Національний університет «Львівська політехніка», Львів, 2024.
11. Комар М. П. «Методологічні основи інформаційної технології інтелектуального аналізу та обробки великих даних». Дисертація на здобуття наукового ступеня доктора технічних наук, 2021. 363 с.
12. Костюченко А. О. Основи програмування мовою Python: навчальний посібник. Ч.: ФОП Баликіна С.М., 2020. 180 с.

Шкала оцінювання:

За шкалою ЗУНУ	За національною шкалою	За шкалою ECTS
90-100	відмінно	A (відмінно)
85-89	добре	B (дуже добре)
75-84		C (добре)
65-74	задовільно	D (задовільно)
60-64		E (достатньо)
35-59	незадовільно	FX (незадовільно з можливістю повторного складання)
1-34		F (незадовільно з обов’язковим повторним курсом)