

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЗАХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ФАКУЛЬТЕТ КОМП'ЮТЕРНИХ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ**

ЗАТВЕРДЖУЮ

**Декан факультету комп'ютерних
інформаційних технологій**


Ігор ЯКИМЕНКО
"29" _____ 2025 р.



ЗАТВЕРДЖУЮ

Проректор з науково-педагогічної роботи


Віктор ОСТРОВЕРХОВ
"29" _____ 2025 р.



**РОБОЧА ПРОГРАМА
з дисципліни**

«Інструментальні засоби обробки даних»

Ступінь вищої освіти – бакалавр

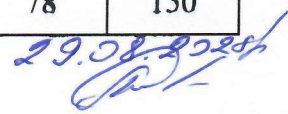
галузь знань – 12 Інформаційні технології

спеціальність – 122 Комп'ютерні науки

Освітньо-професійна програма – «Штучний інтелект»

Кафедра інформаційно-обчислювальних систем і управління

Форма навчання	Курс	Семест р	Лекції (год.)	Лабораторні заняття (год.)	ІРС (год.)	Тренінг (год.)	СРС (год.)	Разом (год.)
Денна	3	6	30	30	4	8	78	150

29.08.2024


**Тернопіль – ЗУНУ
2025**

Робочу програму склав викладач кафедри інформаційно-обчислювальних систем і управління Владислав ПОЙДИЧ.

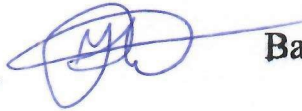
Робоча програма затверджена на засіданні кафедри інформаційно-обчислювальних систем і управління, протокол №1 від 26 серпня 2025 р.

В.о. завідувача
кафедри



Надія ВАСИЛЬКІВ

Гарант освітньо-професійної
програми “ Штучний інтелект ”
к.т.н., доцент



Василь КОВАЛЬ

СТРУКТУРА РОБОЧОЇ ПРОГРАМИ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ «ІНСТРУМЕНТАЛЬНІ ЗАСОБИ ОБРОБКИ ДАНИХ»

1. ОПИС ДИСЦИПЛІНИ

«ІНСТРУМЕНТАЛЬНІ ЗАСОБИ ОБРОБКИ ДАНИХ»

Дисципліна «Інструментальні засоби обробки даних»	Галузь знань, спеціальність, СВО	Характеристика навчальної дисципліни
Кількість кредитів ECTS – 5	Галузь знань – 12 «Інформаційні технології»	Статус дисципліни: вибіркова дисципліна циклу професійної підготовки Мова навчання: Українська
Кількість залікових модулів – 4	Спеціальність – 122 «Комп'ютерні науки»	Рік підготовки: 3 Семестр: <i>Денна – 6</i>
Кількість змістових модулів – 3	Освітньо-професійна програма: Штучний інтелект	Лекції: <i>Денна – 30 год.</i> Лабораторні заняття: <i>Денна – 30 год</i>
Загальна кількість годин – 150	Ступінь вищої освіти – бакалавр	Самостійна робота: <i>Денна – 78 год.,</i> Тренінг: <i>Денна – 8 год.</i> Індивідуальна робота: <i>Денна – 4 год.</i>
Тижневих годин – 10 год., з них аудиторних – 4 год.		Вид підсумкового контролю – залік

2. МЕТА Й ЗАВДАННЯ ДИСЦИПЛІНИ «ІНСТРУМЕНТАЛЬНІ ЗАСОБИ ОБРОБКИ ДАНИХ»

2.1. Мета 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.

2.2. Завдання дисципліни:

- вивчити фундаментальні концепції та принципи, які лежать в основі обробки і аналізу даних, з оглядом на сучасні методи та підходи;
- розвивати практичні навички в області використання сучасних інструментів, мов програмування, платформ і технологій для ефективної обробки даних;
- навчити використовувати технології та інструменти для роботи з великими даними;
- навчити застосовувати сучасні методи та засоби аналітики даних.

2.3. Результати навчання:

В результаті вивчення дисципліни студенти повинні:

Знати:

- інструментальні засоби розробки клієнт-серверних застосувань
- концептуальні, логічні та фізичні моделі баз даних
- оптимізації запитів до бази даних
- розподілені бази даних, сховища та вітрини даних, бази знань, у тому числі на хмарних сервісах, мови веб-програмування

Вміти:

- Застосовувати методи та алгоритми обчислювального інтелекту та інтелектуального аналізу даних в задачах класифікації
- Застосовувати методи та алгоритми прогнозування, кластерного аналізу, пошуку асоціативних правил з використанням програмних інструментів
- Використовувати інструменти підтримки багатовимірного аналізу даних на основі технологій DataMining, TextMining, WebMining.

3. ПРОГРАМА ДИСЦИПЛІНИ «ІНСТРУМЕНТАЛЬНІ ЗАСОБИ ОБРОБКИ ДАНИХ»

ЗМІСТОВИЙ МОДУЛЬ 1. Аналіз та обробка даних.

Тема 1. Вступ в аналіз і обробку даних.

Актуальність даних в сучасному світі. Поняття великих даних. Характеристики великих даних. Методи аналізу та обробки великих даних.

Тема 2. Запити за даними, фільтрування.

Взаємодія з даними. Особливості отримання результатів та їх обробки. Способи фільтрації. Статичні властивості відомих систем, які потрібно виключати при фільтрації.

Тема 3. Об'єднання таблиць: теорія та практика. Особливості типів даних у системах баз даних.

Способи консолідації даних. Типізація об'єднань. Деталізація типів даних, їх взаємодія.

Тема 4. Маніпуляція даними, групування елементів.

Робота з даними на нижчому рівні. Групування елементів. Алгоритми обробки груп та способи ефективного відображення згрупованої інформації.

Тема 5. Транзакції та підзапити. Ефективність дотримання цілісності бази даних.

Поняття транзакції. Види транзакцій. Способи отримання консистентних даних за різних умов навантаження серверів. Робота з матеріалізацією інформації у форматі підзапитів.

Тема 6. Обмеження для забезпечення цілісності та консистентності збережених даних.

Аналіз видів обмежень для забезпечення правильності структури бази даних, цілісності та послідовності. Ефективність обмежень на рівні таблиці та рівні рядка. Зовнішні обмеження.

ЗМІСТОВИЙ МОДУЛЬ 2. Технології та інструменти роботи з великими даними.

Тема 7. Бази даних та сховища даних.

Особливості баз даних NoSQL. Організація та принципи функціонування MongoDB, PostgreSQL. Дослідження швидкодії роботи різних баз даних.

Тема 8. Розподілені платформи.

Інфраструктура розподілених обчислень Hadoop. Інфраструктура розподілених обчислень Spark. Інфраструктура розподілених обчислень Kafka.

Тема 9. Інструментальні засоби.

Вимоги до програмних систем. Система DBT. Система Apache Airflow. Система Snowflake. Система Fivetran. Мови програмування: R, Python. DBT - SQL та Jinja.

ЗМІСТОВИЙ МОДУЛЬ 3. Сучасні методи та засоби аналітики даних. Візуалізація даних.

Тема 10. Платформи обробки даних.

Особливості архітектури DBT. Робота з деталями процесу перетворення. Способи реалізації снєпшотів.

Тема 11. Методи запуску циклічних задач по оброці даних Apache Airflow.

Що таке циклічні задачі. Способи оптимізації процесів перетворення за допомогою Apache Airflow. Мови реалізації DAG файлів.

Тема 12. Методи перетворення даних у процесі ETL та ELT, DBT.

Основні відмінності та застосування: ETL (Extract, Transform, Load) та ELT (Extract, Load, Transform). Переваги використання DBT як інструмента для перетворення даних. Тестування результатів. Забезпечення цілісності.

Тема 13. Засоби візуалізації чистих даних: PowerBI, Looker.

Що таке візуалізація даних. Порівняльна характеристика систем PowerBI та Looker. Термінологія візуалізації, Stories, Dataset, Datasource, Data Platform.

Тема 14. PowerBI запити, підключення до зовнішніх джерел даних, перетворення та особливості агрегування даних.

Погляд всередину роботи з PowerBI. Отримання даних з зовнішнього джерела. Перетворення інформації. Різниця між перетворенням даних у DBT та перетворенням даних у PowerBI.

4. СТРУКТУРА ЗАЛІКОВОГО КРЕДИТУ ДИСЦИПЛІНИ «ІНСТРУМЕНТАЛЬНІ ЗАСОБИ ОБРОБКИ ДАНИХ»

Денна форма навчання

2	Кількість годин					
	Лек.	Лабор	СРС	ІР	Тренінг	Контрольні заходи
ЗМІСТОВИЙ МОДУЛЬ 1. Аналіз та обробка даних.						
Тема 1. Вступ в аналіз і обробку даних.	2	-	4	2	4	Опитування під час заняття
Тема 2. Запити за даними, фільтрування.	2	2	4			Опитування під час заняття
Тема 3. Об'єднання таблиць: теорія та практика. Особливості типів даних у системах баз даних.	2	2	6			Опитування під час заняття
Тема 4. Маніпуляція даними, групування елементів.	2	2	6			Опитування під час заняття
Тема 5. Транзакції та підзапити. Ефективність дотримання цілісності бази даних.	4	2	10			Опитування під час заняття
Тема 6. Обмеження для забезпечення цілісності та консистентності збережених даних.	2	2	6			Опитування під час заняття
ЗМІСТОВИЙ МОДУЛЬ 2. Технології та інструменти роботи з великими даними						
Тема 7. Бази даних та сховища даних.	2	-	4	1	2	Опитування під час заняття
Тема 8. Розподілені платформи.	2	-	4			Опитування під час заняття
Тема 9. Інструментальні засоби.	2	2	6			Опитування під час заняття
ЗМІСТОВИЙ МОДУЛЬ 3. Сучасні методи та засоби аналітики даних. Візуалізація даних.						
Тема 10. Платформи обробки даних.	2	2	6	1	2	Опитування під час заняття
Тема 11. Методи запуску циклічних задач по обробці даних Apache Airflow.	2	2	6			Опитування під час заняття
Тема 12. Методи перетворення даних у процесі ETL та ELT, DBT.	2	4	6			Опитування під час заняття
Тема 13. Засоби візуалізації чистих даних: PowerBI, Looker.	2	6	4			Опитування під час заняття
Тема 14. PowerBI запити, підключення до зовнішніх джерел даних, перетворення та особливості агрегування даних.	2	4	6			Опитування під час заняття
Всього	30	30	78	4	8	

5. ТЕМАТИКА ЛАБОРАТОРНИХ ЗАНЯТЬ

Лабораторна робота №1.

Тема: Робота з даними.

Мета: Освоїти методи взаємодії з даними через мову SQL.

Лабораторна робота №2.

Тема: Транзакційність та підзапити.

Мета: Навчитися створювати транзакції та будувати матеріалізовані підзапити.

Лабораторна робота №3.

Тема: Apache Airflow.

Мета: Освоїти Apache Airflow для автоматизації процесів ETL/ELT.

Лабораторна робота №4.

Тема: DBT.

Мета: Освоїти DBT для імплементації процесу перетворення даних з використанням різних типів сховищ.

Лабораторна робота №5.

Тема: DBT тести та обмеження.

Мета: Освоїти деталізацію процесів перетворення та забезпечити цілісність результатів за допомогою перевірки перетворених даних.

Лабораторна робота №6.

Тема: PowerBI та візуалізація.

Мета: Навчитися працювати з візуалізатором PowerBI, використати дані попередніх лабораторних занять для візуалізації.

6. САМОСТІЙНА РОБОТА

Самостійно опрацювати наступні тем:

- Основні характеристики великих даних
- Роль великих даних в різних галузях
- Консолідація даних
- Мова запитів SQL
- SQL транзакції та підзапити
- Архітектура Snowflake
- Архітектура Hadoop
- Виконання Map/Reduce
- Робота з Apache Airflow
- Створення DAG
- DBT база
- DBT перетворення
- DBT тестування
- Підключення DBT та Airflow для автоматизації
- PowerBI концепт
- PowerBI візуалізація
- PowerBI PowerQuery
- Looker як спосіб візуалізації
- Порівняння Looker та PowerBI
- Використання датасетів різного типу у PowerBI.

Реферат.

Написати реферат на тему «Порівняльна характеристика розподілених платформ Hadoop, Kafka, Apache Spark». Розмір реферату не менше 8 сторінок. Включити порівняльну таблицю платформ.

Практичне завдання.

Створити структуру даних у постгресі. Налаштувати DBT перетворення. Налаштувати Apache Airflow Dag для циклічного запуску процесів перетворення DBT у даних з структури у постгресі.

Умовами успішного виконання є захищений реферат, та захист практичного завдання.

7. ТРЕНІНГ З ДИСЦИПЛІНИ

Реалізувати комплексний проєкт. Комплексний проєкт повинен складатися з наступних частин:

1. Потсгрес база даних, що містить структуру даних довільної сфери.
2. Трансформація даних за допомогою ETL/ELT процесів.
3. Візуалізація даних за допомогою сервісу PowerBI.

Результатом виконання проєкту повинен бути створений Application у PowerBI сервісі який візуалізує потрібні дані з структуру за допомогою:

1. Кругова діаграма
2. Гістограма
3. Кластерна гістограма

8. ЗАСОБИ ОЦІНЮВАННЯ ТА МЕТОДИ ДЕМОНСТРУВАННЯ РЕЗУЛЬТАТІВ НАВЧАННЯ

У навчальному процесі застосовуються: лекції, в тому числі з використанням мультимедійного проектора та інших ТЗН, практичні заняття, консультації, індивідуальна та самостійна робота студента.

В процесі вивчення дисципліни «Інструментальні засоби обробки даних» використовуються наступні методи оцінювання навчальної роботи студента:

- поточне опитування;
- оцінювання лабораторних робіт;
- модульна контрольна робота;
- оцінювання виконання завдання самостійної роботи;
- оцінювання виконання завдань під час тренінгу.

9. КРИТЕРІЇ, ФОРМИ ПОТОЧНОГО ТА ПІДСУМКОВОГО КОНТРОЛЮ

Підсумковий бал (за 100-бальною шкалою) з дисципліни «Інструментальні засоби обробки даних» визначається як середньозважена величина, в залежності від питомої ваги кожної складової залікового кредиту:

Модуль 1		Модуль 2		Модуль 3	Модуль 4
20%	20%	20%	20%	5%	15%
Поточне оцінювання	Модульний контроль	Поточне оцінювання	Модульний контроль	Тренінг	Самостійна робота
Середнє арифметичне за лабораторні роботи та опитування на заняттях.	Тестовий контроль у системі Moodle 20 тестів по 2 бали Дві задачі по 15 балів Теоретичне питання 30 балів	Середнє арифметичне за лабораторні роботи та опитування на заняттях.	Тестовий контроль у системі Moodle 20 тестів по 2 бали Дві задачі по 15 балів Теоретичне питання 30 балів	Виконання завдання комплексного проєкту тренінгу.	Виконання завдання самостійної роботи.

Модуль 1 та Модуль 2 включають дві складові за якими оцінюють знання студента:

1. **“Поточне оцінювання”** - виставляється як середнє арифметичне за лабораторні роботи та опитування на заняттях.

ECTS, бали	Критерії оцінювання
A 90 – 100	Виконані всі завдання лабораторної роботи. Реалізований код програми є працездатним, оптимізованим та структурованим. Правильно виконані всі етапи роботи (налаштування середовища, програмування, відлагодження, перевірка результатів). Звіт оформлено акуратно з поясненням принципів роботи програми. Студент впевнено відповідає на запитання викладача, надає повні пояснення, вміє узагальнити результати та відповідає на додаткові запитання. Звіт оформлено акуратно з поясненням принципів роботи програм та без помилок.
B 85 – 89	Виконані всі основні завдання та частина додаткових. Результати працездатні, хоча присутні дрібні неточності у кодї чи описі. Студент демонструє добре розуміння теми та пояснює ключові рішення. Звіт містить незначні недоліки в оформленні.
C 75 – 84	Виконані всі основні завдання, але без додаткових. Програма працює, але з обмеженнями (відсутні додаткові перевірки, код програми не оптимізований, відсутня обробка виняткових ситуацій). Частково зроблений аналіз отриманих результатів. У звіті є недоліки в оформленні або опис.
D	Виконана лише частина основних завдань. Представлені рішення містять суттєві недоліки в кодї або логіці роботи, але студент демонструє базове

ECTS, бали	Критерії оцінювання
65 – 74	розуміння теми. Робота виконана за зразком, без власних пояснень. Аналіз та висновки неповні або неточні. У звіті є помилки чи пропуски.
E 60 – 64	Виконано мінімальну частину базових завдань з допомогою викладача. Отримані результати обмежено працездатні. Аналіз та висновки відсутні або поверхневі. Звіт має значні недоліки.
FX 35 – 59	Більшість завдань не виконано. Код програми або відсутній, або непрацездатний. Студент не може пояснити свої дії та не розуміє суті лабораторної роботи. Звіт відсутній або не відповідає вимогам.

2. **Модульний контроль**” – проводиться за допомогою тестового контролю по пройдених темах у системі Moodle. 20 тестів по 2 бали Дві задачі по 15 балів. Теоретичне питання 30 балів. Кожен тест – це набір питань де студент повинен вибрати правильну відповідь. Тести формуються у вигляді питань так/ні, вибрати правильні серед багатьох питань.

Теоретичне питання оцінюються за наступними критеріями:

- 20-30 – відповідь надана повністю і чітко, інформація подана явно.
- 10-20 – інформація відповідає питанню, але частково і основна тема не розкрита.
- 0-10 – інформація не відповідає питанню, основна тема частково освітлена.

Задачі оцінюються за наступними критеріями:

- 10-15 - Правильно виконана задача
- 8-10. Частково виконана задача, правильний напрямок виконання
- 1-8. Частково виконана задача, неправильний напрямок виконання
- 0. Не виконана задача

Модуль 3 - “Тренінг” – оцінка визначається, як середнє арифметичне оцінок, отриманих за виконання окремих завдань тренінгу, відповідно п. 7 даної робочої програми.

ECTS, бали	Критерії оцінювання
A 90 – 100	Завдання тренінгу виконані в повному обсязі та реалізовані на високому рівні з урахуванням всіх вимог. Показана креативність або оптимізація, що покращує результат. Студент упевнено та аргументовано пояснює прийняті рішення, відповідає на запитання. Демонструє впевнене володіння матеріалом. Продемонстровано якісний кінцевий програмний продукт/прототип.
B 85 – 89	Виконано всі основні завдання. Присутні незначні недоліки у деталях або оформленні. Результат відповідає вимогам та є цілком працездатним.
C 75 – 84	Виконано більшість завдань. Деякі вимоги реалізовані поверхово. Представлений результат працездатний, але має обмежену функціональність.
D 65 – 74	Завдання виконано частково. Присутні суттєві недоліки в реалізації або якості результату. Досягнуто мінімально необхідного рівня. Реалізована основна ідея, але без глибокої деталізації.
E 60 – 64	Виконано лише базову частину завдань. Робота відповідає мінімальним вимогам та демонструє обмежений рівень знань і практичних умінь.
FX 35 – 59	Завдання виконано частково, значна частина ключових етапів відсутня. Результат неповний або не відповідає встановленим вимогам. Демонстрація роботи програми/прототипу обмежена або неможлива.

Модуль 4 - “Самостійна робота” – оцінка визначається, як середнє арифметичне оцінок, отриманих під час виконання завдання самостійної роботи, а саме: 1) представлення проведеного дослідження з обраної теми; 2) аналіз отриманих результатів.

ECTS, бали	Критерії оцінювання
A 90 – 100	Робота виконана повністю, тема розкрита всебічно. Використано достатню кількість джерел, є чітка структура (вступ, основна частина, висновки). Аналіз результатів глибокий та аргументований, зроблено власні висновки. Оформлення відповідає вимогам.
B 85 – 89	Тема досліджена на високому рівні, використано якісні джерела. Присутня логічна структура та аналіз, проте глибина висновків або критичний огляд обмежені. Невеликі недоліки в оформленні.
C 75 – 84	Виконано основний обсяг роботи, тема розкрита частково. Аналіз поверховий, висновки загальні. Використано обмежену кількість джерел, але робота

ECTS, бали	Критерії оцінювання
	залишається цілісною та зрозумілою.
D 65 – 74	Тема досліджена неповністю, значна частина матеріалу подана описово без належного аналізу. Висновки слабкі або нечіткі. Є помітні недоліки в структурі та оформленні.
E 60 – 64	Представлено лише базову інформацію по темі, відсутній достатній аналіз. Висновки формальні. Робота відповідає мінімальним вимогам.
FX 35 – 59	Виконано частину завдання. Тема розкрита поверхово, аналіз практично відсутній. Структура нечітка, джерела використані обмежено. Робота не відповідає більшості вимог

Шкала оцінювання:

За шкалою Університету	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)

10. ІНСТРУМЕНТИ, ОБЛАДНАННЯ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ, ВИКОРИСТАННЯ ЯКИХ ПЕРЕДБАЧАЄ НАВЧАЛЬНА ДИСЦИПЛІНА

№	Найменування	Номер теми
1.	Мультимедійне обладнання	1-14
2.	Комп'ютери з доступом до мережі Інтернет, високопродуктивний сервер	2-14
3.	Технології та інструменти роботи з даними	2-14

РЕКОМЕНДОВАНІ ДЖЕРЕЛА ІНФОРМАЦІЇ

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. Microsoft Power BI Dashboards Step by Step, Errin O'Connor, 2018.
4. PostgreSQL: Up and Running: A Practical Guide to the Advanced Open Source Database, Regina Obe, Leo Hsu, 2017.
5. Python for Data Analysis: Data Wrangling with pandas, NumPy, and Jupyter, 3rd Edition, Wes McKinney, 2022.
6. Zgurovsky M.Z., Zaychenko Y.P. Big Data: Conceptual Analysis and Applications. - Springer, 2020. – 298 p.
7. Wiktorski Tomasz. Data-intensive Systems: Principles and Fundamentals using Hadoop and Spark. - Springer, 2019. – 105 p.
8. Bühlmann P., Drineas P., Kane M., van der Laan M. (eds.) Handbook of Big Data. -Taylor and Francis, 2016. – 456 p.
9. Oussous A., Benjelloun F.-Z., Lahcen A.A. and Belfkih S. Big Data technologies: A survey. // Journal of King Saud University. Computer and Information Sciences. - 2018. - Vol. 30, Issue 4. - P. 431–448.
10. Akerkar R. Models of Computation for Big Data Cham: Springer International Publishing, 2018. – 110 p.
11. Raheem N. Big Data: A Tutorial-Based Approach. - Taylor & Francis Group LLC, CRC Press, 2019. – 203 p.
12. Комар М.П. «Методологічні основи інформаційної технології інтелектуального аналізу та обробки великих даних». Дисертація на здобуття наукового ступеня доктора технічних наук, 2021. – 363 с.
13. Костюченко А.О. Основи програмування мовою Python: навчальний посібник. - Ч.: ФОП Баликіна С.М., 2020. - 180 с.
14. Опорний конспект лекцій з дисципліни «Інструментальні засоби обробки даних» для студентів спеціальності F3 «Комп'ютерні науки» першого (бакалаврського) рівня вищої освіти. Укладач Пойдич В.С. Тернопіль: ЗУНУ, 2025. 59 с.