

Міністерство освіти і науки України
Національний університет «Запорізька політехніка»

кафедра програмних засобів

ЗВІТ
з лабораторної роботи № 1
з дисципліни «Інтелектуальний аналіз даних» на тему:
**«РОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ІНТЕЛЕКТУАЛЬНОГО
АНАЛІЗУ ДАНИХ»**

Виконав:

ст. гр. КНТ-113сп

Іван ЩЕДРОВСЬКИЙ

Прийняв:

зав. лабораторії

Максим АНДРЕЄВ

2025

1 Мета роботи

Ознайомитися та отримати навички роботи з графічним інтерфейсом користувача бібліотеки data mining алгоритмів WEKA та мовою програмування Python. Вивчити основні можливості програми WEKA та мови програмування Python

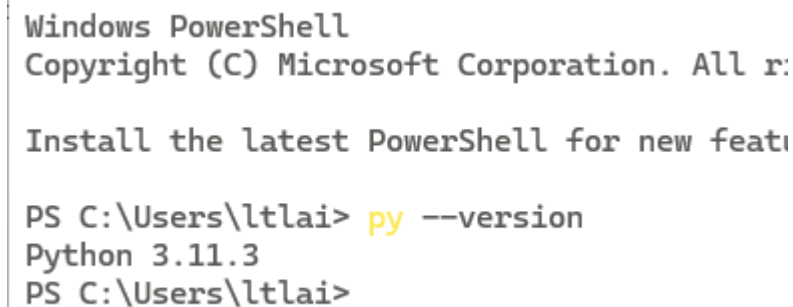
2 Завдання до лабораторної роботи

2.1 Встановити програму для інтелектуального аналізу даних WEKA та інтерпретатор мови програмування Python

2.2 Ознайомитися з основними можливостями графічного інтерфейсу користувача програми WEKA та основами програмування на мові Python

3 Виконання лабораторної роботи

Було встановлено Python версії 3.11.3, це показано на рисунку 3.1



```
Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

Install the latest PowerShell for new features and optimizations!

PS C:\Users\ltlai> py --version
Python 3.11.3
PS C:\Users\ltlai>
```

Рисунок 3.1 – Версія встановленого Python

Також для виконання робіт було встановлено IDE PyCharm від JetBrains. Ця IDE призначена для розробки саме на Python та має багато вбудованих систем та налаштувань. Як альтернатива можна було б налаштувати neovim з нуля, але це займе більше часу та буде складніше для початку роботи з аналізом даних

На рисунку 3.2 показано інтерфейс PyCharm

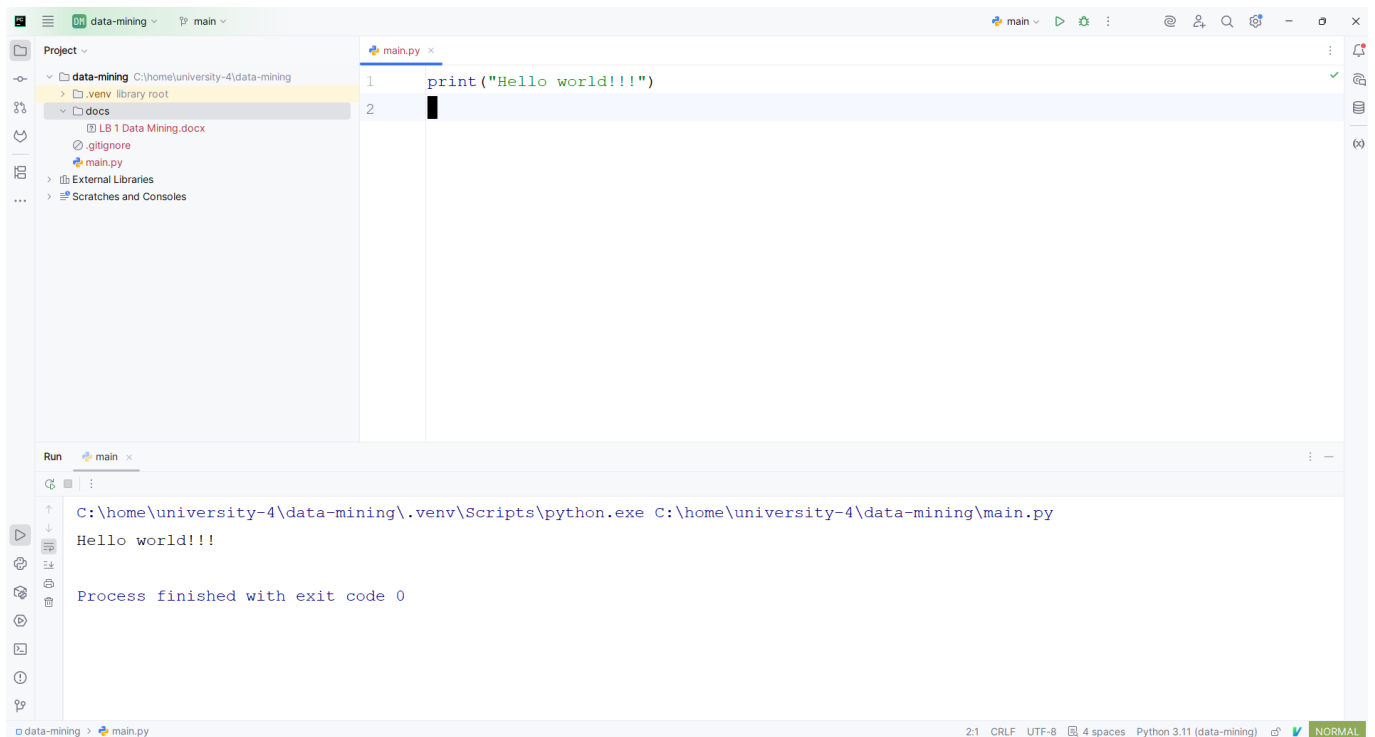


Рисунок 3.2 – Інтерфейс PyCharm

Щодо ознайомлення, то з мовою Python я був знайомий раніше та маю декілька проєктів де вона використовується

Далі було встановлено Weka версії 3.9.6. Інтерфейс цієї програми показаний на рисунку 3.3



Рисунок 3.3 – Інтерфейс програми Weka

Було переглянуто її інтерфейс по інструкції з методичних вказівок

4 Відповіді на контрольні запитання

Що таке інтелектуальний аналіз даних?

Це сукупність методів для знаходження раніше не відомих, неявних практично корисних та доступних для інтерпретації знань, які можуть використовуватись в різних сферах життя

Які основні задачі вирішує інтелектуальний аналіз даних?

Основна задача знаходити зв'язки та шаблони в великому наборі даних. Наприклад, знайти які товари люди частіше за все купують разом з якимось іншим товаром. Це може використовуватись для оптимізації маршруту людини в магазині

Що таке задача класифікації?

У нас є задана скінчена кількість об'єктів, вибірка, для яких відомо до яких класів кожен з них належить

Також є інші об'єкти, для яких потрібно цей клас визначити

Задача класифікації це і є визначення класу для об'єкта

Важливо зауважити, що в цьому завданні є обмежена кількість класів, а також ці класи можуть задаватись оператором, або ж братись із даних.

Що таке задача регресії?

Методи аналізу залежності однієї величини від іншої

На відміну від кореляційного аналізу не з'ясовує чи істотний зв'язок, а займається пошуком моделі цього зв'язку, вираженої у функції регресії

Що таке задача кластеризації?

Задача розбиття вибірки об'єктів на підмножини, які називаються кластерами

Кожен кластер повинен складатись зі схожих об'єктів, а об'єкти з різних кластерів повинні суттєво відрізнятись

Що таке задача пошуку асоціативних правил?

Метод машинного навчання заснованого на правилах для знаходження цікавих відношень між змінними у великих базах даних.

Метою є ідентифікація сильних правил, які виявляються в базах даних з використанням деяких вимірів зацікавленості

Які основні властивості мови програмування Python?

Я не до кінця розумію що мається на увазі під «властивостями»

Мова Python є мультипарадигменою, високорівневою, популярною, інтерпритованою

Які парадигми програмування підтримує мова Python?

Об'єктно-орієнтоване, функціональне та структурне програмування

Які типи даних підтримує мова Python?

Текст, числа, а саме int, float та complex. Списки, tuple та range. Словники(асоціативний масив), множини(Set), bool, тощо

Які команди pip ви знаєте?

help, install, uninstall, search, check, show, list, freeze

Як встановити модуль чи пакет?

pip install

Що необхідно зробити, щоб використати встановлену бібліотеку в коді програми за допомогою мови Python?

Її потрібно імпортувати, або ж можна імпортувати частково якийсь метод бібліотеки. Також деякі бібліотеки підключаються через декілька імпортів

Опишіть формат файлів даних, який буде використовуватися при роботі з Python.

В Python для даних буде використовуватись формат csv, в розшифровці comma-separated values

Це формат файлу який представляє собою таблицю, де дані, або ж різні стовпці, розділені комами. Перший рядок представляє собою заголовки для таблиці

Якщо якісь значення відсутні то зазвичай їх пропускають, або можуть писати спеціальні символи або текст, наприклад ? або NaN, тощо

Також в цьому форматі може використовуватись інший роздільник замість коми

5 Висновки

Я ознайомився та отримав навички роботи з графічним інтерфейсом користувача бібліотеки data mining алгоритмів WEKA та мовою програмування Python. Вивчив основні можливості програми WEKA та мови програмування Python