**Лабораторна робота №2**

**Робота з візуалізацією даних та Streamlit.**

**Мета роботи:** набуття навичок роботи із фреймворком Streamlit.

**Хід роботи:**

Для підготовки до роботи слід використовувати матеріали з лекцій та результати першої лабораторної роботи.

Для роботи з файлами \*.ipynb рекомендується встановити Anaconda Navigator за посиланням <https://www.anaconda.com/distribution/> і запустіть JupiterLab. Лабораторні роботи рекомендується виконувати в середовищі JupiterLab (JupiterNotebook).

Для роботи із файлами \*.py можна використовувати Spider, що також є в пакеті Anaconda Navigator.

Перед роботою із Streamlit необхідно його встановити за допомогою команди

*ʼpip install streamlitʼ* будь яким зручним для вас способом.

1. В середовищі Spider (чи іншому зручному для вас) створити файл lab2.py, імпортувати необхідні для виконання дослідження бібліотеки, а також прописати

import streamlit as st

В подальшому для запуску файлу прописати в терміналі

streamlit run lab2.py

Для розміщення елементів в додатку, що розробляється можна використовувати будь які підходи (експандери, кнопки, колонки і т.ін.), що подробно описані в <https://docs.streamlit.io/library/api-reference>

2. Завантажити в dataframe ту саму вибірку даних, що і в першій лабораторній роботі (табл.1). Тут може виникнути 2 опції:

- якщо PyCaret та Streamlit встановлені в різні conda environment, то доведеться зберегти файл на диск з PyCaret, щоб потім використати його в Streamlit;

- якщо PyCaret та Streamlit встановлені в той самий conda environment, то можна завантажити безпосередньо з PyCaret.

Таблиця 1

|  |  |  |
| --- | --- | --- |
| № | Прізвище студента | Назва датасету |
|  | Біліхін | nba |
|  | Гріцай | bike |
|  | Кавецькайте | hepatitis |
|  | Канцемал | airquality |
|  | Краснюк | diamond |
|  | Левчишин | country-data |
|  | Леонов | wine |
|  | Масалітіна | pokemon |
|  | Мітрошкіна | forest |
|  | Мусієнко | bank |
|  | Рожко | income |
|  | Селезньова | boston |
|  | Соколов | titanic |
|  | Струнов | automobile |
|  | Тристан | glass |
|  | Шапаренко | churn |
|  | РЕЗЕРВ | us\_presidential\_election\_results |

Для кожного датасету знайти в інтернеті опис ознак для повного розуміння фізичного сенсу даних, що ви маєте. Додати цей опис за допомогою, наприклад, st.write() чи st.table().

2. В залежності від завантаженого датасету необхідно провести візуалізацію даних та відобразити результати. Варіантами того, що ви можете зробити є гістограми, стовпчикові та кругові діаграми з накопиченням та без, боксплоти та ін. Вибір способу візуалізації безпосередньо залежить від датасету. Графіки мають містити назву, позначки на осях, бажано використовувати елементи інтерактивності (наприклад, відображати частину даних використовуючи st.multiselect(), st.radio(), st.checkbox() тощо).

3. Використати результати отримані при виконанні першої лабораторної роботи, а саме pkl файл із навченою моделлю та відобразити результати навчання у вигляді метрик, графіків тощо.

**Звіт**

Звітом є файл із кодом програми у форматі \*.py який містить streamlit компоненти.