

## Звіт

з дисципліни «Проектування кібер-фізичних систем»  
на тему «Передбачення погоди»

студента групи ТВ-33  
Домарацького Дмитра

### 5 варіант

#### Завдання:

5) Аналітика по якості повітря - візуалізація (коли можна алергикам виходити, коли берегтись, як змінюється в залежності від сезону, чи є зв'язок з іншими параметрами погоди). перший датасет.

#### Розв'язання:

Планую реалізувати на мові Python у Visual Studio.

Будемо користуватись, за вказівкою у самому завданні, наданим датасетом з сайту Kaggle. Для початку я завантажив csv-файл за посиланням <https://www.kaggle.com/datasets/nelgiriyeewithana/global-weather-repository?resource=download> і розмістив його в папці проекту (рисунок 1).

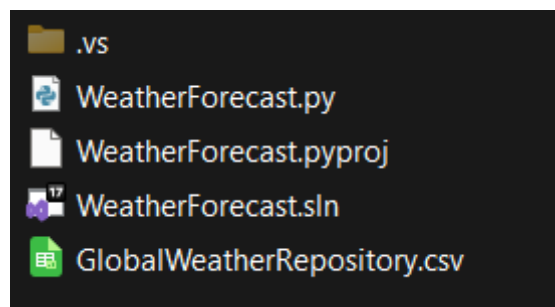


Рисунок 1 — розміщення датасету

Далі мені потрібно ознайомитись із теоретичною частиною та зрозуміти що використовувати для досягнення цілі: машинне навчання чи алгоритм, якщо і машинне навчання, то за яким алгоритмом навчати.

Думаю, що виконаю завдання не застосовуючи машинне навчання, оскільки мені потрібна аналітика якості повітря. Далі проаналізуємо csv-файл (рисунок 2, 3).

- **country:** Country of the weather data
- **location\_name:** Name of the location (city)
- **latitude:** Latitude coordinate of the location
- **longitude:** Longitude coordinate of the location
- **timezone:** Timezone of the location
- **last\_updated\_epoch:** Unix timestamp of the last data update
- **last\_updated:** Local time of the last data update

Рисунок 2 — деякі корисні колонки

- **air\_quality\_Carbon\_Monoxide:** Air quality measurement: Carbon Monoxide
- **air\_quality\_Ozone:** Air quality measurement: Ozone
- **air\_quality\_Nitrogen\_dioxide:** Air quality measurement: Nitrogen Dioxide
- **air\_quality\_Sulphur\_dioxide:** Air quality measurement: Sulphur Dioxide
- **air\_quality\_PM2.5:** Air quality measurement: PM2.5
- **air\_quality\_PM10:** Air quality measurement: PM10
- **air\_quality\_us-epa-index:** Air quality measurement: US EPA Index
- **air\_quality\_gb-defra-index:** Air quality measurement: GB DEFRA Index

Рисунок 3 — основні колонки

Тут я виділив основні, на мою думку, колонки з датасету, якими я збираюсь користуватись для аналізу даних. Також про інші забувати не буду, оскільки, згідно з завданням, потрібно зрозуміти чи є кореляція між якістю повітрям і іншими параметрами погоди, такими як: температура, тиск, кількість опадів і т.д. Також варто зауважити, що збираюсь робити аналіз тільки по Україні, оскільки з прикладної точки зору це має більше сенсу (принаймні так мені здається).

Дані потрібно аналізувати та візуалізувати, для цього імпортуємо стандартні бібліотеки для такого типу задач: numpy, pandas, matplotlib.pyplot (рисунок 4).

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd

data = pd.read_csv("GlobalWeatherRepository.csv")
```

Рисунок 4 — імпортування бібліотек та зчитування файлу

Тепер давайте побудуємо графіки (рисунок 5).

```
def draw():
    plt.legend()
    plt.grid()
    plt.show()

plt.plot(ua['last_updated'][::AVERAGE_STEP], ua['air_quality_Carbon_Monoxide'])
plt.plot(ua['last_updated'][::AVERAGE_STEP], ua['air_quality_Ozone'])
draw()

plt.plot(ua['last_updated'][::AVERAGE_STEP], ua['air_quality_Nitrogen_dioxide'])
plt.plot(ua['last_updated'][::AVERAGE_STEP], ua['air_quality_Sulphur_dioxide'])
draw()

plt.plot(ua['last_updated'][::AVERAGE_STEP], ua['air_quality_PM2.5'])
plt.plot(ua['last_updated'][::AVERAGE_STEP], ua['air_quality_PM10'])
draw()

plt.plot(ua['last_updated'][::AVERAGE_STEP], ua['air_quality_us-epa-index'])
plt.plot(ua['last_updated'][::AVERAGE_STEP], ua['air_quality_gb-defra-index'])
draw()
```

Рисунок 5 — код для малювання графіків

В кінці зрозумів, що бібліотека `numpy` не використовується, тому прибираю її.

За основу візьмемо документацію ВООЗ останньої редакції для встановлення допустимих норм забруднюючих речовин в повітрі: <https://www.who.int/publications/i/item/9789240034228>. На сторінці 135 є таблицка 3.24 з відповідними нормами усіх основних забруднюючих речовин.

Тепер коли є графіки та допустимі норми, можна проаналізувати результат (рисунок 6, 7, 8, 9, 10). В датасеті є дані з минулого травня по теперішній.

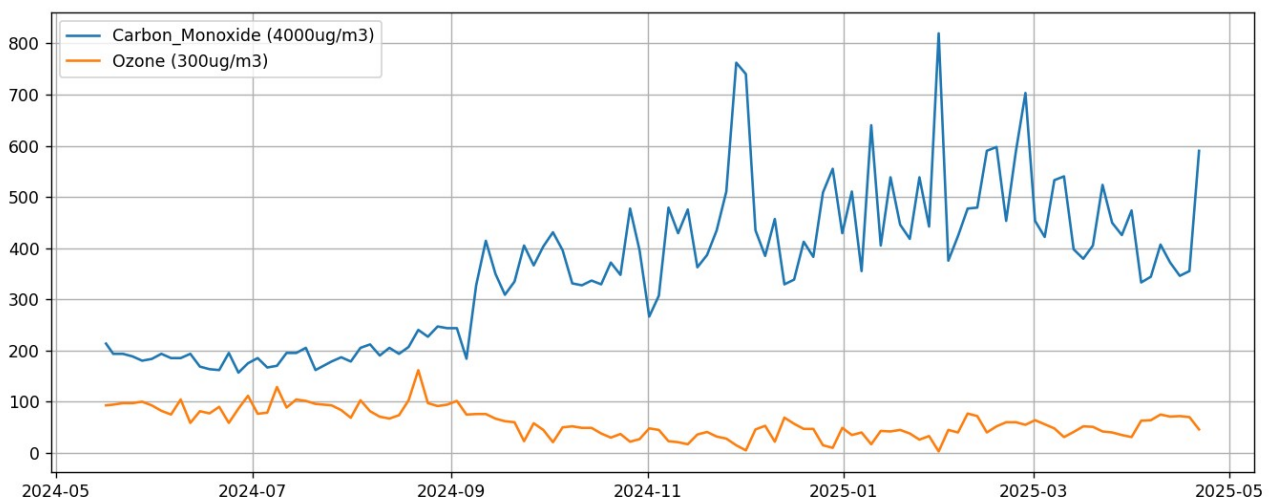


Рисунок 6 — графіки чадного газу та озону

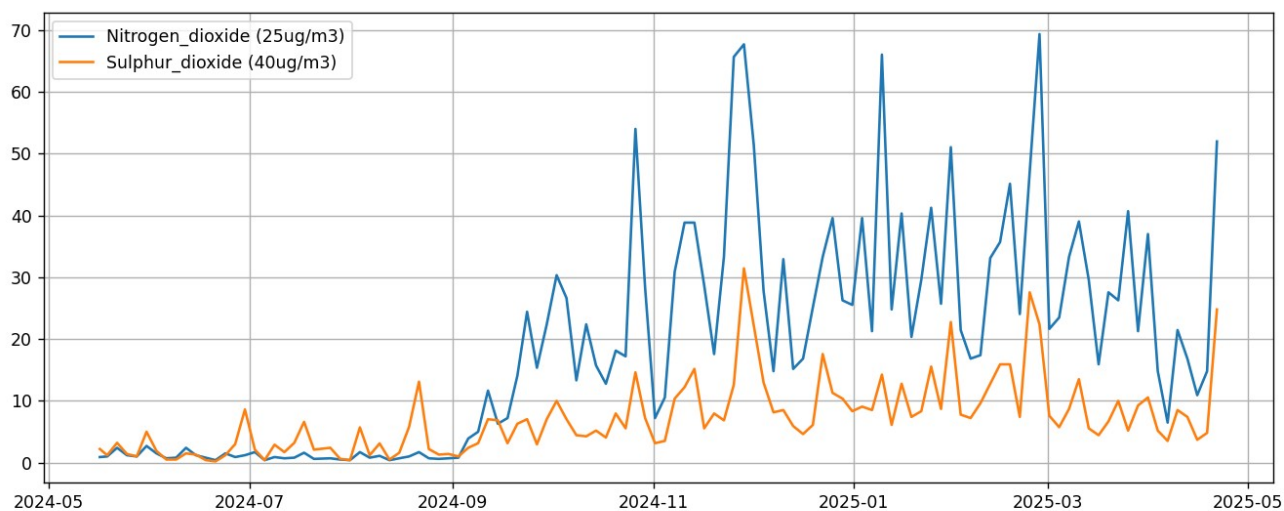


Рисунок 7 — графіки діоксиду азоту та сірки

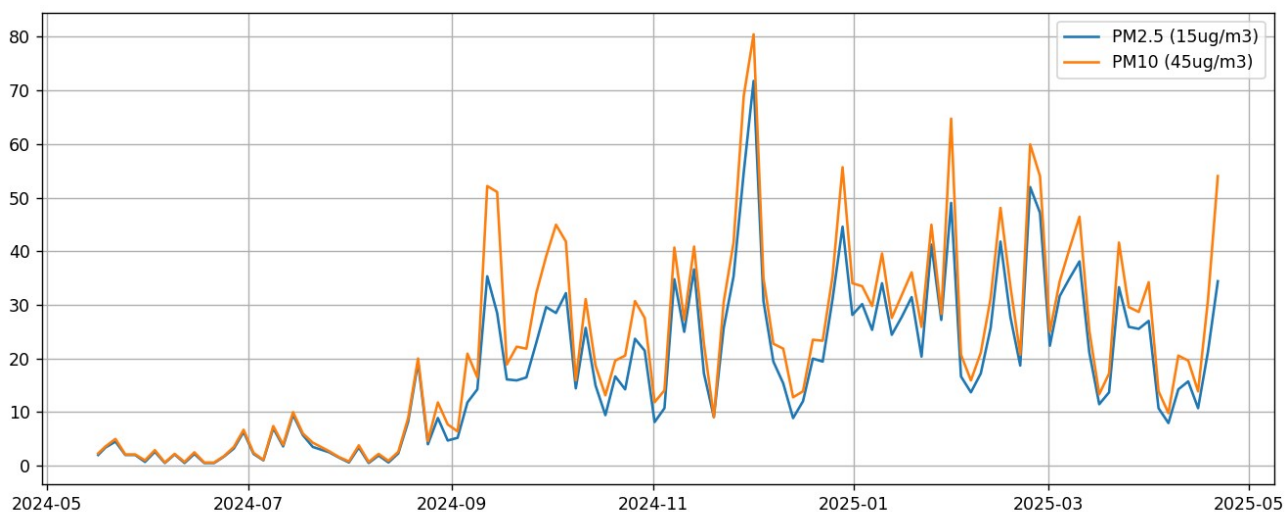


Рисунок 8 — графіки забруднюючих частинок різного розміру

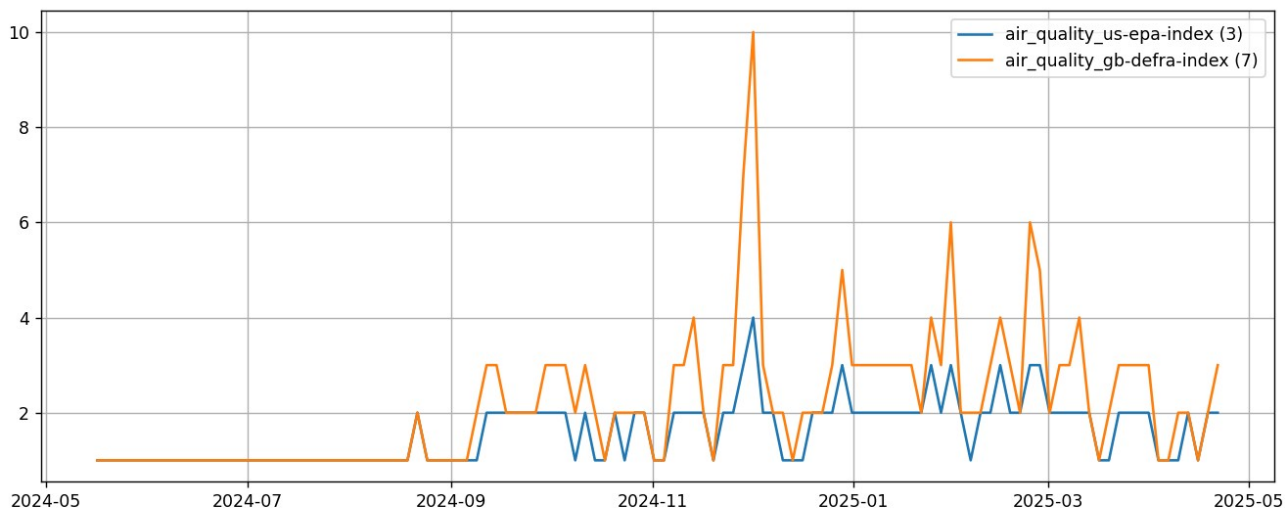


Рисунок 9 — графіки британських та американських індексів

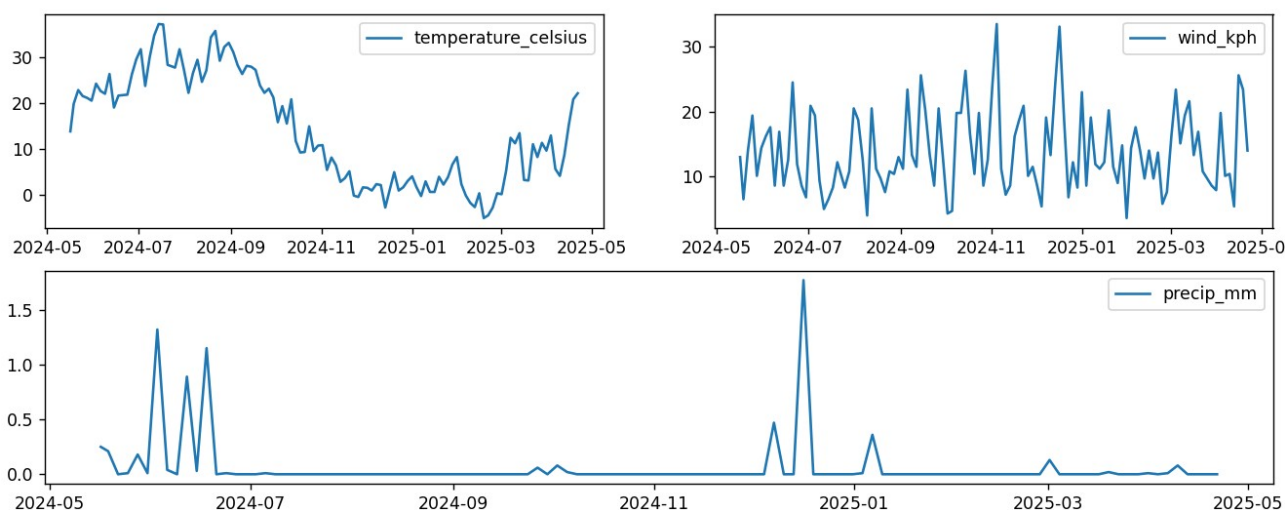


Рисунок 10 — графіки інших параметрів погоди

З графіків можна помітити наступну тенденцію: кількість забруднюючих речовин (окрім озону) мінімальна наприкінці весни та протягом усього літа, збільшується з приходом осені та зими, перевищуючи допустиму норму, і починає зменшуватись всередині цієї весни.