

Міністерство освіти і науки України
НТУУ «КПІ ім. Ігоря Сікорського»
Навчально-науковий інститут атомної та теплової енергетики
Кафедра цифрових технологій в енергетиці

Лабораторна робота №8
з дисципліни «Вступ до інтелектуального аналізу даних»
Тема «Візуалізація часових рядів»
Варіант №19

Студента 3-го курсу НН ІАТЕ гр. ТР-12

Ковальова Олександра

Перевірив: д.т.н., проф. Путренко В. В.

Мета: Опрацювати приклад роботи з візуалізацією часових рядів, використовуючи Jupyter Notebook. Виконати поставлене завдання.

Хід роботи

Для початку підключаємо потрібні бібліотеки:

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

Executed at 2024.04.14 17:35:44 in 3ms

Для аналізу мінімальних денних температур взятих зі спостережень, було використано функцію `np.genfromtxt`, що дозволяє зчитати дані з файлу у форматі CSV. Ця функція створює масив даних, в якому кожен рядок відповідає окремому запису, а значення в кожному рядку розділені вказаним роздільником.

Параметр `delimiter` вказує символ-роздільник, який використовується в CSV файлі. В даному випадку, в якості роздільника використовується кома. Параметр `skip_header` дозволяє пропустити задану кількість рядків зверху файлу, у цьому випадку було пропущено перший рядок, оскільки він містив заголовок.

Після завантаження даних, колонка, що містить мінімальні температури, була виділена в окремий масив за допомогою індексації.

Для обчислення середнього значення та стандартного відхилення використовувалися відповідно функції `np.nanmean()` та `np.nanstd()`. Ці функції використовуються для обчислення середнього значення та стандартного відхилення відповідно, ігноруючи значення NaN у вхідних даних.

```
1 data = np.genfromtxt('./data/daily-minimum-temperatures-in-me.csv',
2                       delimiter = ",", skip_header = 1)
3 temps = data[:,1]
4 mean = np.nanmean(temps)
5 std = np.nanstd(temps)
```

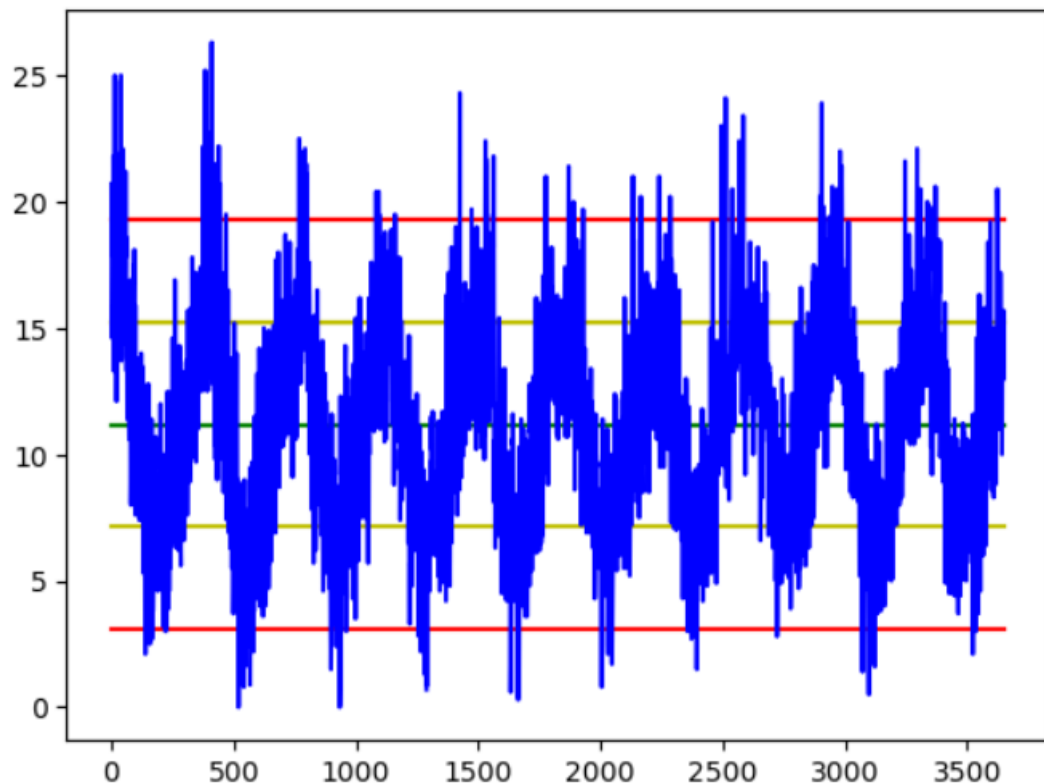
Executed at 2024.04.14 17:35:44 in 13ms

При візуалізації даних було використано бібліотеку `matplotlib.pyplot` для побудови графіку. Кожна з команд `plt.plot()` відповідає за побудову однієї лінії на графіку.

```
1 plt.plot([0,len(temps)], [mean, mean], 'g-')
2 plt.plot([0,len(temps)], [mean+std, mean+std], 'y-')
3 plt.plot([0,len(temps)], [mean, mean], 'g-')
4 plt.plot([0,len(temps)], [mean+std, mean+std], 'y-')
5 plt.plot([0,len(temps)], [mean-std, mean-std], 'y-')
6 plt.plot([0,len(temps)], [mean+2*std, mean+2*std], 'r-')
7 plt.plot([0,len(temps)], [mean-2*std, mean-2*std], 'r-')
8 plt.plot([0,len(temps)], [mean, mean], 'g-')
9 plt.plot([0,len(temps)], [mean+std, mean+std], 'y-')
10 plt.plot([0,len(temps)], [mean-std, mean-std], 'y-')
11 plt.plot([0,len(temps)], [mean+2*std, mean+2*std], 'r-')
12 plt.plot([0,len(temps)], [mean-2*std, mean-2*std], 'r-')
13 plt.plot(range(len(temps)), temps, 'b-')
14 plt.show()
```

Executed at 2024.04.14 17:35:44 in 105ms

Результат:



Висновок: Під час виконання лабораторної роботи були набуті практичні навички роботи з візуалізацією часових рядів. Було виконане поставлене завдання.

Програмний код

Notebook.ipynb:

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

data = np.genfromtxt('./data/daily-minimum-temperatures-in-me.csv',
                    delimiter = ",", skip_header = 1)

temps = data[:,1]
mean = np.nanmean(temps)
std = np.nanstd(temps)

plt.plot([0,len(temps)], [mean, mean], 'g-')
plt.plot([0,len(temps)], [mean+std, mean+std], 'y-')
plt.plot([0,len(temps)], [mean, mean], 'g-')
plt.plot([0,len(temps)], [mean+std,mean+std], 'y-')
plt.plot([0,len(temps)], [mean-std, mean-std], 'y-')
plt.plot([0,len(temps)], [mean+2*std, mean+2*std], 'r-')
plt.plot([0,len(temps)], [mean-2*std, mean-2*std], 'r-')
plt.plot([0,len(temps)], [mean, mean], 'g-')
plt.plot([0,len(temps)], [mean+std, mean+std], 'y-')
plt.plot([0,len(temps)], [mean-std, mean-std], 'y-')
plt.plot([0,len(temps)], [mean+2*std, mean+2*std], 'r-')
plt.plot([0,len(temps)], [mean-2*std, mean-2*std], 'r-')
plt.plot(range(len(temps)), temps, 'b-')
plt.show()
```