Національний технічний університет України «КПІ ім. Ігоря Сікорського»

Факультет Інформатики та Обчислювальної Техніки

Кафедра інформаційних систем та технологій

Лабораторна робота №2

з дисципліни «Прикладні задачі машинного навчання»

на тему

«Часові ряди і проста лінійна регресія»

15 варіант

Виконав:

студент групи ІС-21 Костюк А. С.

Викладач:

Нестерук А.О.

Київ – 2024

**Зміст**

[**1. Постановка задачі** 3](#_Toc162295224)

[**2. Виконання** 4](#_Toc162295225)

[**3. Висновок** 12](#_Toc162295226)

**1. Постановка задачі**

1. В даній лабораторній роботі Вам треба завантажити метеорологічні дані в 1895-2022 роках з CSV-файлу в DataFrame. Після цього дані треба буде відформатувати для використання.

2. Бібліотеку Seaborn використати для графічного представлення даних DataFrame у вигляді регресійної прямої, що представляє графік зміни обраних показників за період 1895-2018 років.

3. Спрогнозуйте дані на 2019, 2020, 2021 та 2022, 2023 рік.

4. Оцініть за формулою, якою могли б бути показники до 1895 року.

5. Скористайтесь функцією regplot бібліотеки Seaborn для виведення всіх точок даних; дати представляються на осі x, а показники на осі y. Функція regplot будує діаграму розкиду даних, на якій точки представляють показники за заданий рік, а пряма лінія - регресійну пряму.

6. Виконайте масштабування осі у від (приклад від 10 до 70 градусів).

7. Порівняйте отриманий прогноз для 2019, 2020, 2021 та за 2022, 2023 роки з даними на NOAA «Climate at a Glance»: https://www.ncdc.noaa.gov/cag/ і зробити висновок.

8. Зробити звіт про роботу.

**2. Виконання**

1) **В даній лабораторній роботі Вам треба завантажити метеорологічні дані в 1895-2023 роках з CSV-файлу в DataFrame. Після цього дані треба буде відформатувати для використання.**

Для дослідження обрав місто Spokane, що знаходиться в штаті Washington. Метеорологічні дані: середня температура в березні по роках.

Зображення, що містить текст, знімок екрана, Графік, ряд

Автоматично згенерований опис

Зображення, що містить текст, знімок екрана, програмне забезпечення

Автоматично згенерований опис

Зображення, що містить текст, знімок екрана, Шрифт, число

Автоматично згенерований опис

Зображення, що містить текст, знімок екрана, програмне забезпечення

Автоматично згенерований опис

Зображення, що містить текст, знімок екрана, Шрифт

Автоматично згенерований опис

2) **Бібліотеку Seaborn використати для графічного представлення даних DataFrame у вигляді регресійної прямої, що представляє графік зміни обраних показників за період 1895-2018 років.**

Зображення, що містить текст, знімок екрана, програмне забезпечення, Мультимедійне програмне забезпечення

Автоматично згенерований опис

Зображення, що містить знімок екрана, текст, монітор, програмне забезпечення

Автоматично згенерований опис

3) **Спрогнозуйте дані на 2019, 2020, 2021 та 2022, 2023 рік.**

Зображення, що містить текст, знімок екрана, Шрифт, програмне забезпечення

Автоматично згенерований опис

Зображення, що містить текст, знімок екрана, Шрифт

Автоматично згенерований опис

4) **Оцініть за формулою, якою могли б бути показники до 1895 року.**

Зображення, що містить текст, знімок екрана, Шрифт

Автоматично згенерований опис

5) **Скористайтесь функцією regplot бібліотеки Seaborn для виведення всіх точок даних; дати представляються на осі x, а показники на осі y. Функція regplot будує діаграму розкиду даних, на якій точки представляють показники за заданий рік, а пряма лінія - регресійну пряму.**

Зображення, що містить знімок екрана, текст, монітор, програмне забезпечення

Автоматично згенерований опис

6) **Виконайте масштабування осі у від (приклад від 10 до 70 градусів).**

Зображення, що містить текст, знімок екрана, програмне забезпечення

Автоматично згенерований опис

7) **Порівняйте отриманий прогноз для 2019, 2020, 2021 та за 2022, 2023 роки з даними на NOAA «Climate at a Glance»: https://www.ncdc.noaa.gov/cag/ і зробити висновок.**

Зображення, що містить текст, знімок екрана, програмне забезпечення

Автоматично згенерований опис

Отриманий результат вказує на те, що шляхом прогнозування погоди методом порівняння статистики погоди за минулі роки не вдалося досягти точних результатів. Причиною цього є хибність припущення про лінійну зміну погоди, оскільки погода має складну та непередбачувану характеристику. Наприклад, зміна напрямку вітру, температури, вологості повітря та різні антропогенні фактори можуть спричинити зміну погодних умов незалежно від статистики погоди за минулі роки.

**3. Висновок**

В даній лабораторній роботі були розглянуті часові ряди та проста лінійна регресія, обробка та візуалізація яких була здійснена за допомогою Python та його бібліотек SciPy, statistics, numpy, seaborn та pandas. Отримані знання були використані для обробки даних середньої температури в березні 1895-2023 років в місті Spokane штату Washington, а результати представлені в лабораторній.