

**ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH**  
**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**  
**KHOA CÔNG NGHỆ THÔNG TIN**



**BÁO CÁO**  
**ĐỒ ÁN THỰC HÀNH CUỐI KỲ**  
**MÔN: TOÁN ỨNG DỤNG VÀ THỐNG KÊ**

**Lớp: 21\_2**

Giảng viên hướng dẫn: **ThS. Nguyễn Hữu Toàn**  
**Cô Võ Nam Thục Đoan**

Sinh viên thực hiện: **Lê Trần Minh Khuê**

MSSV: **21120279**

THÀNH PHỐ HỒ CHÍ MINH, THÁNG 6/2023

## Lời cảm ơn

Em chân thành cảm ơn thầy Nguyễn Hữu Toàn và cô Võ Nam Thực Doan, hai giảng viên đã dày công truyền đạt kiến thức và hướng dẫn em trong suốt môn học này.

Em đã cố gắng vận dụng những kiến thức đã học được để hoàn thành đồ án. Nhưng do kiến thức hạn chế và chưa có nhiều kinh nghiệm nên khó tránh khỏi những thiếu sót trong quá trình nghiên cứu và trình bày.

Rất kính mong sự góp ý của quý thầy cô để bài báo cáo của em được hoàn thiện hơn.

Em xin trân trọng cảm ơn sự quan tâm giúp đỡ của cô và thầy trong suốt môn học này.

Xin trân trọng cảm ơn!

# Mục lục

<b>1</b>	<b>Đề bài</b>	<b>3</b>
1.1	Đề 03: Xích Markov . . . . .	3
1.2	Chọn địa điểm: . . . . .	3
<b>2</b>	<b>Tổng quan về Xích Markov</b>	<b>3</b>
2.1	Định nghĩa và khái niệm cơ bản. . . . .	3
2.2	Ma trận chuyển trạng thái. . . . .	3
2.3	Vector phân phối xác suất tại thời điểm $t$ . . . . .	4
2.4	Các công thức: . . . . .	4
2.5	Phân phối dừng . . . . .	4
<b>3</b>	<b>Mô hình Xích Markov cho dự báo thời tiết</b>	<b>5</b>
<b>4</b>	<b>Thu thập và xử lý dữ liệu thời tiết</b>	<b>5</b>
4.1	Xác định các trạng thái thời tiết . . . . .	5
4.2	Thu thập dữ liệu: . . . . .	6
4.3	Quy ước biểu tượng và thời tiết tương ứng . . . . .	6
<b>5</b>	<b>Lập ma trận chuyển trạng thái</b>	<b>6</b>
<b>6</b>	<b>Lập trình mô hình dự báo thời tiết bằng Xích Markov</b>	<b>7</b>
6.1	Các hàm hỗ trợ . . . . .	7
6.2	Các hàm tính toán Xích Markov: . . . . .	10
6.3	Các hàm nhập dữ liệu: . . . . .	10
6.4	Phần chương trình: . . . . .	11
6.4.1	Với bộ dữ liệu gán sẵn . . . . .	11
6.4.2	Với bộ dữ liệu nhập vào . . . . .	11
6.4.3	Các phần dự đoán: . . . . .	11
<b>7</b>	<b>Phụ lục</b>	<b>11</b>
7.1	Định dạng file chứa ma trận chuyển trạng thái . . . . .	11
7.2	Định dạng file chứa ma trận tần suất chuyển trạng thái . . . . .	12
7.3	Thư mục lưu các file dữ liệu mẫu: . . . . .	12
<b>8</b>	<b>Tài liệu tham khảo</b>	<b>12</b>

# 1 Đề bài

## 1.1 Đề 03: Xích Markov

Ứng dụng xích Markov vào 1 bài toán cụ thể trong thực tế qua việc dự báo thời tiết ngày hôm sau dựa trên thời tiết hiện tại.

## 1.2 Chọn địa điểm:

Em lựa chọn Thành phố Thủ Đức - Thành phố Hồ Chí Minh là địa điểm để thu thập dữ liệu và tiến hành xây dựng mô hình dự báo thời tiết cho địa điểm này.

# 2 Tổng quan về Xích Markov

## 2.1 Định nghĩa và khái niệm cơ bản.

Xích Markov là một mô hình xác suất được sử dụng để mô phỏng và dự đoán các quá trình ngẫu nhiên theo thời gian. Đặc trưng quan trọng của Xích Markov là tính Markov, tức là trạng thái tương lai chỉ phụ thuộc vào trạng thái hiện tại và không phụ thuộc vào quá khứ.

- Giả sử một hệ thống vật lý hoặc toán học trải qua một quá trình thay đổi sao cho tại bất kỳ thời điểm nào, nó có thể có một trạng thái trong *một số hữu hạn trạng thái* (ký hiệu là  $S = \{s_1, s_2, \dots, s_m\}$ ).
- Giả sử rằng một hệ thống như vậy thay đổi theo thời gian từ trạng thái này sang trạng thái khác và tại các thời điểm xác định, trạng thái của hệ thống được quan sát.
- Nếu trạng thái của hệ thống tại bất kỳ quan sát nào không thể được dự đoán một cách chắc chắn, nhưng *xác suất xảy ra của một trạng thái nhất định có thể được dự đoán chỉ bằng cách biết trạng thái của hệ thống tại lần quan sát trước đó*, thì quá trình thay đổi được gọi là xích Markov hoặc quy trình Markov.

## 2.2 Ma trận chuyển trạng thái.

Ma trận chuyển trạng thái biểu diễn xác suất chuyển từ một trạng thái sang trạng thái khác trong một bước.

Nếu một xích Markov có  $k$  trạng thái, mà chúng ta đặt tên là  $1, 2, \dots, k$ , thì xác suất hệ thống ở trạng thái  $i$  tại thời điểm ngay sau khi nó ở trạng thái  $j$  tại thời điểm trước đó được ký hiệu là  $P_{ij}$  và được gọi là xác suất chuyển từ trạng thái  $j$  sang trạng thái  $i$ . Ma trận  $P = [p_{ij}]$  được gọi là ma trận chuyển tiếp của xích Markov.

**Ví dụ một ma trận chuyển 3 trạng thái của xích Markov:**

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

Trong đó,  $p_{12}$  là xác suất chuyển từ trạng thái 2 sang trạng thái 1,  $p_{31}$  là xác suất chuyển từ trạng thái 1 sang trạng thái 3,...

**Các tính chất quan trọng của ma trận chuyển trạng thái:**

- Xác suất chuyển từ một trạng thái sang tất cả các trạng thái khác là không âm và tổng các xác suất chuyển phải bằng 1.

$$p_{1j} + p_{2j} + \dots + p_{kj} = 1$$

- Các xác suất chuyển trong ma trận xác suất chuyển có thể thay đổi theo thời gian (nếu gặp yếu tố bên ngoài tác động) hoặc thường được coi là không thay đổi.

## 2.3 Vector phân phối xác suất tại thời điểm t.

Vector  $\pi_n = (\pi_n(1), \pi_n(2), \dots, \pi_n(m))$  trong đó  $\pi_n(i) = P(X_n = s_i)$  được gọi là phân phối của xích tại thời điểm  $t = n$ . Phân phối của  $X_0$ ,  $\pi_0$  được gọi là phân phối đầu của xích.

## 2.4 Các công thức:

- a) Với mọi  $n$ , ta có  $P(X_{t+n} = s_i | X_t = s_j) = P(X_n = s_i | X_0 = s_j)$ ,  $\forall i, j = 1, 2, \dots, m$

Với mỗi  $n$ , đặt  $p_{ij}^n = P(X_n = s_i | X_0 = s_j)$ ,  $\forall i, j = 1, 2, \dots, m$

Ma trận  $P^{(n)} = [p_{ij}^n]$  được gọi là *ma trận chuyển trạng thái sau n bước*.

- b) Với mọi  $n$ ,  $P^{(n)} = P^n$ , trong đó  $P^n$  là lũy thừa  $n$  của ma trận chuyển  $P$ , ta có

$$\pi_{t+n} = P^n \pi_t$$

$$\pi_n = P^n \pi_0$$

## 2.5 Phân phối dừng

Là một phân phối xác suất mà khi đến một thời điểm nào đó hệ thống đạt đến trạng thái ổn định, phân phối xác suất này không thay đổi theo thời gian nữa.

Phân phối  $\pi = (\pi(1), \pi(2), \dots, \pi(m))$  thỏa  $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \pi(i)$  với mọi  $i, j = 1, 2, \dots, m$  thì phân phối  $\pi$  được gọi là phân phối giới hạn của xích.

**Phân phối  $\pi$  thỏa  $\pi = P\pi$  được gọi là phân phối dừng của xích.**

- Xích có thể không có, có đúng một hoặc nhiều phân phối dừng<sup>(\*)</sup>.

- Khi xích rơi vào một phân phối dừng thì xích sẽ dừng ở phân phối đó (xích không thay đổi phân phối), nghĩa là nếu có  $n$  để  $\pi_n = \pi$  là một phân phối dừng thì  $\pi_t = \pi$  với mọi  $t \geq n$ .
- Phân phối giới hạn nếu có, là một phân phối dừng.

**Ma trận chính quy:** Ma trận  $P$  được gọi là chính quy nếu tồn tại  $n \in \mathbb{N}$  sao cho mọi số hạng của  $P^n$  đều là số dương.

(\*): Nếu xích Markov có ma trận chuyển  $P$  là chính quy thì phân phối giới hạn cũng chính là phân phối dừng duy nhất.

### 3 Mô hình Xích Markov cho dự báo thời tiết

Trong bài toán dự báo thời tiết, Xích Markov có thể được áp dụng để mô hình hóa sự chuyển đổi giữa các trạng thái thời tiết và dự đoán trạng thái thời tiết trong tương lai dựa trên thông tin thời tiết hiện tại. Quá trình dự báo thời tiết bằng Xích Markov được thực hiện theo các bước sau:

1. Thu thập dữ liệu thời tiết: Thu thập các thông tin thời tiết hiện tại từ các nguồn dữ liệu như trạm quan trắc thời tiết hoặc các trang web dự báo thời tiết.
2. Xác định các trạng thái thời tiết: Xác định các trạng thái thời tiết có thể có, chẳng hạn như mưa, nắng, giông, mây,...
3. Xây dựng ma trận chuyển trạng thái: Dựa trên dữ liệu thời tiết đã thu thập, xác định xác suất chuyển từ mỗi trạng thái thời tiết hiện tại sang các trạng thái thời tiết khác.
4. Dự báo thời tiết: Sử dụng ma trận chuyển trạng thái và trạng thái thời tiết hiện tại, áp dụng công thức Xích Markov để tính toán xác suất của các trạng thái thời tiết trong tương lai.
5. Đánh giá dự báo: So sánh kết quả dự báo với dữ liệu thực tế để đánh giá hiệu suất của mô hình dự báo. Điều này giúp cải thiện và điều chỉnh mô hình Xích Markov để đạt được dự báo thời tiết chính xác hơn. (Trong phạm vi đồ án này, em tập trung hơn về phần xử lý dữ liệu và dự đoán, chưa có hướng tiếp cận để đánh giá kết quả, nếu có thì chỉ có thể đánh giá trên một số lượng dự báo nhất định, chưa thể phát triển nhiều.)

### 4 Thu thập và xử lý dữ liệu thời tiết

#### 4.1 Xác định các trạng thái thời tiết

Trước khi áp dụng Xích Markov vào dự báo thời tiết, ta cần xác định các trạng thái thời tiết có thể có.

4 trạng thái phổ biến nhất ở Thành phố Thủ Đức - TPHCM là:

- Nắng/ Sunny: Trời trong và có nắng nhiều.
- Nhiều mây/ Cloudy: Trời nhiều mây nhưng không mưa, có thể có nắng nhẹ.
- Mưa/ Rainy: Trời có mưa.
- Giông bão/ Stormy: Trời mưa lớn đến rất lớn / Bão.

## 4.2 Thu thập dữ liệu:

Sau quá trình tìm kiếm và xem xét giữa nhiều nguồn, nhiều website dự báo thời tiết trong và ngoài nước, em quyết định lựa chọn website "[The Weather Channel](https://www.weather.com)" vì miễn phí và có lưu lại dữ liệu thời tiết trước đó,...

Đường dẫn truy cập vào dữ liệu thời tiết ở khu vực Thành phố Thủ Đức: [Dự báo thời tiết hàng tháng Thủ Đức - weather.com](https://www.weather.com)

Dữ liệu thời tiết được thu thập trong 1 năm (365 ngày) từ ngày 18/06/2022 đến ngày 17/06/2023.

## 4.3 Quy ước biểu tượng và thời tiết tương ứng

- Nắng/ Sunny: Biểu tượng mặt trời không mây và mặt trời bị che bởi ít mây.
- Nhiều mây/ Cloudy: Biểu tượng mây và mặt trời bị che bởi nhiều mây.
- Mưa/ Rainy: Biểu tượng mây + mưa - không sấm sét.
- Giông bão/ Stormy: Biểu tượng mây + mưa - có sấm sét.

Ngày	Nắng	Nhiều mây	Mưa	Giông bão
6/18/2022		x		
6/19/2022		x		
6/20/2022			x	
6/21/2022				x
6/22/2022			x	

Bảng 1: Mẫu dữ liệu minh họa

## 5 Lập ma trận chuyển trạng thái

Để xây dựng ma trận chuyển trạng thái, cần tính toán và ước lượng xác suất chuyển từ dữ liệu thời tiết đã thu thập.

Em sử dụng tần suất xuất hiện của các trạng thái thời tiết và các cặp trạng thái liên kề để ước lượng xác suất chuyển:

- Từ dữ liệu thời tiết, ta đếm số lần mỗi trạng thái thời tiết xuất hiện và số lần mỗi cặp trạng thái xuất hiện liên kề. (\*Lưu ý: Dữ liệu được lấy từ 365 ngày, do đó, số lần

chuyển trạng thái là 364 lần)

	Nắng	Nhiều mây	Mưa	Giông bão
Nắng	38	36	4	0
Nhiều mây	34	134	30	10
Mưa	6	26	16	7
Giông bão	0	13	5	5
TOTAL	78	209	55	22

Bảng 2: Bảng thống kê tần suất xuất hiện của mỗi trạng thái và các cặp trạng thái

- o Sau đó, ta chia số lần xuất hiện của mỗi cặp trạng thái cho số lần xuất hiện của trạng thái thời tiết hiện tại để tính toán xác suất chuyển (Làm tròn đến 5 chữ số thập phân).

STATISTIC	Nắng	Nhiều mây	Mưa	Giông bão
Nắng	0.48718	0.17225	0.07273	0
Nhiều mây	0.43590	0.64115	0.54545	0.45455
Mưa	0.07692	0.12440	0.29091	0.31818
Giông bão	0	0.06220	0.09091	0.22727
TOTAL	1	1	1	1

Bảng 3: Bảng xác suất chuyển trạng thái

Vậy ta thu được ma trận chuyển trạng thái  $P$  là:

$$P = \begin{bmatrix} 0.48718 & 0.17225 & 0.07273 & 0 \\ 0.43590 & 0.64115 & 0.54545 & 0.45455 \\ 0.07692 & 0.12440 & 0.29091 & 0.31818 \\ 0 & 0.06220 & 0.09091 & 0.22727 \end{bmatrix}$$

## 6 Lập trình mô hình dự báo thời tiết bằng Xích Markov

### 6.1 Các hàm hỗ trợ

Trước khi nhập dữ liệu (ma trận chuyển trạng thái -  $P$ ) và xây dựng các hàm cho việc dự báo thời tiết, em sẽ tiến hành xây dựng trước các hàm hỗ trợ cần thiết.

Lưu ý: Vì yêu cầu không được gọi thêm bất cứ thư viện nào khác, nên em sẽ cài đặt lại hết tất cả các hàm cần thiết.

Danh sách các hàm hỗ trợ như sau:

1. **is\_zero(x)**: Hàm kiểm tra số rất gần 0.



- Hàm nhận vào tham số x là số cần kiểm tra.
- Hàm trả về kết quả True hoặc False.

### Các hàm hỗ trợ trong việc tính toán trên vector:

2. **CopyV(vector)**: Hàm copy vector, khi biến đổi vector mới sẽ không làm ảnh hưởng vector ban đầu.
3. **is\_Zero\_Vector(vec)**: Hàm kiểm tra tất cả các phần tử của vector là 0
4. **multiply\_vectors(vector1, vector2)**: Hàm nhân từng phần tử của hai vector đầu vào. Trong trường hợp hai vector không cùng kích thước, thông báo lỗi và trả về **None**.
5. **multiply\_scalar\_vector(vector, scalar)**: Hàm nhân vector với hằng số scalar.
6. **add\_vectors(vectorA, vectorB)**: Hàm cộng hai vector. Trong trường hợp hai vector không cùng kích thước, thông báo lỗi và trả về **None**.

### Các hàm hỗ trợ trong việc tính toán trên ma trận:

7. **print\_matrix(matrix)**: Hàm in ma trận (dùng cho ma trận dạng list với mỗi phần tử cũng là list).
8. **identity\_matrix(n)**: Hàm tạo ma trận đơn vị cấp n.
9. **get\_column(matrix, i)**: Hàm lấy cột thứ i của ma trận.
  - Trong trường hợp i lớn hơn số cột của ma trận sẽ thông báo lỗi và trả về **None**.
  - Ngược lại, trả về vector cột thứ i của ma trận.
10. **CopyM(matrix)**: Hàm sao chép ma trận. Khi xử lý trên ma trận mới sẽ không làm ảnh hưởng đến ma trận ban đầu.
11. **rows\_permutation(matrix, row1, row2)**: Hàm hoán vị 2 dòng của ma trận với các tham số đầu vào là ma trận và chỉ số của 2 dòng cần hoán vị.
  - Trường hợp có chỉ số vượt quá số dòng của ma trận sẽ thông báo lỗi và trả về **None**.
  - Ngược lại, trả về ma trận với 2 dòng row1 và row2 đã được hoán vị.
12. **multiply\_scalar\_matrix(matrix, scalar)**: Nhân ma trận matrix với hằng số scalar.
13. **multiply\_matrices(mat1, mat2)**: Nhân hai ma trận mat1 và mat2.
  - Nếu số cột của ma trận mat1 và số dòng của ma trận mat2 không bằng nhau: Thông báo lỗi và trả về **None**.
  - Ngược lại, trả về ma trận kết quả.
14. **add\_matrices(mat1, mat2)**: Hàm cộng hai ma trận.
  - Nếu kích thước của hai ma trận không bằng nhau: Thông báo lỗi và trả về **None**.
  - Ngược lại, trả về ma trận kết quả có các phần tử ở vị trí [i][j] là tổng của mat1[i][j] và mat2[i][j].

15. **power\_matrix(matrix, power)**: Hàm lũy thừa ma trận.
- Nếu  $\text{power} < 0$  hoặc ma trận không vuông: Thông báo lỗi tương ứng và trả về **None**.
  - Ngược lại, trả về ma trận lũy thừa tương ứng.

**Các hàm tính toán giữa ma trận và vector:**

16. **multiply\_matrix\_vector(matrix, vector)**: Hàm nhân ma trận với vector.
- Nếu số phần tử của vector và số cột của ma trận không bằng nhau: Thông báo lỗi và trả về **None**.
  - Ngược lại trả về kết quả phép nhân.
17. **multiply\_vector\_matrix(vector, matrix)**: Hàm nhân vector dòng với ma trận.
- Nếu số phần tử của vector và số dòng của ma trận không bằng nhau: Thông báo lỗi và trả về **None**.
  - Ngược lại trả về kết quả phép nhân.

**Các hàm gọi chung cho ma trận và vector:**

18. **get\_variable\_type(variable)**: Hàm kiểm tra biến variable là ma trận hay vector.
- Xác định biến đưa vào là matrix hay vector (list) bằng hàm `isinstance(object, classinfo)`
- Các kiểu trả về:
- 0: Nếu biến không phải là một list.
  - 1: Nếu biến là một list nhưng không phải là ma trận (một list chứa các phần tử không phải list).
  - 2: Nếu biến là một ma trận (một list chứa các phần tử là list).
19. **Copy(variable)**: gọi hàm `get_variable_type` để kiểm tra kiểu của biến và gọi hàm `Copy` tương ứng.
20. **multiply\_MV(var1, var2)**: gọi hàm `get_variable_type` để kiểm tra kiểu của biến và gọi hàm nhân phù hợp.

**Các hàm tính toán, giải hệ phương trình tuyến tính:** Áp dụng cho việc giải tìm vector phân phối dừng.

21. **create\_augmented\_matrix(matrix, vector)**: Hàm tạo ma trận bổ sung từ matrix và vector.
- Nếu số phần tử của vector và số dòng của ma trận không bằng nhau: Thông báo lỗi và trả về **None**.
  - Ngược lại trả về ma trận bổ sung tương ứng.
22. **Gauss\_Jordan\_elimination(mat)**: Hàm biến đổi mat về dạng bậc thang rút gọn bằng giải thuật Gauss Jordan.
23. **add\_coefficient(matrix, row1, row2, coef)**: Hàm cộng hệ số, xử lý giải hệ phương trình vô số nghiệm.

24. **complete\_coefficient(matrix,no,value,coef)**: Hàm hoàn tất hệ số cho nghiệm, xử lý giải hệ phương trình vô số nghiệm.
25. **back\_substitution(matrix)**: Hàm thế ngược tìm nghiệm.
  - Trường hợp vô nghiệm: Thông báo và trả về **None**.
  - Trường hợp nghiệm duy nhất: Trả về vector nghiệm.
  - Trường hợp vô số nghiệm: Trả về ma trận nghiệm với dòng  $i$  tương ứng với  $x_i$  và cột  $j$  tương ứng với vector hệ số của ẩn tự do (nếu  $j=0$  thì vector tương ứng là các hệ số tự do).

## 6.2 Các hàm tính toán Xích Markov:

1. **forecast\_next\_day(state\_vector,P)**: Hàm dự đoán thời tiết cho ngày liền kề, dựa vào vector phân phối đầu và ma trận chuyển P.
2. **forecast\_xth\_day(state\_vector,P,x)**: Hàm dự đoán thời tiết cho ngày thứ  $x$  kể từ ngày hiện tại, dựa vào vector phân phối đầu và ma trận chuyển P.
3. **print\_forecast(state,forecast\_state)**: Hàm xuất kết quả dự đoán ở dạng tỉ lệ phần trăm xảy ra của trạng thái.
  - Nhận vào vector state chứa tên các trạng thái và forecast\_state chứa vector dự báo.
  - Nếu forecast\_state là vector 0 thì báo lỗi.
  - Ngược lại, in kết quả dự đoán.
4. **is\_Regular(P)**: Hàm kiểm tra ma trận chính quy.
5. **steady\_state\_vector(P)**: Hàm tìm vector phân phối dừng.
  - B1: Kiểm tra tính chính quy của ma trận P.
  - B2: Lập hệ phương trình  $(I_n - P)\pi = 0$
  - B3: Giải tìm vector riêng tương ứng của hệ nghiệm.
  - B4: Tính toán  $\pi$  dựa trên tính chất tổng xác suất  $=1$ .

## 6.3 Các hàm nhập dữ liệu:

1. **get\_counting(path)**: Hàm nhập ma trận đếm sự chuyển trạng thái từ file csv.
2. **cal\_P(path)**: Hàm tính ma trận chuyển trạng thái dựa trên dữ liệu đọc vào. (Gọi hàm get\_counting(path) để lấy dữ liệu).
3. **get\_P(path)**: Nhập ma trận chuyển trạng thái từ file csv.
4. **input\_Weather\_State\_vector()**: Hàm nhập vector phân phối đầu từ bàn phím. Nhập vào 1 trạng thái thời tiết (tiếng Anh). Được phép nhập lại tối đa 2 lần. Nếu vẫn sai, mặc định tạo vector phân phối đầu là vector 0.
5. **input\_transition\_matrix()**: Hàm nhập vào các trạng thái (vector state) và ma trận chuyển trạng thái từ bàn phím.

## 6.4 Phần chương trình:

### 6.4.1 Với bộ dữ liệu gán sẵn

Bộ dữ liệu gán sẵn bao gồm:

- Ma trận chuyển trạng thái P (ở phần 5).
- Vector lưu tên các trạng thái `state = ['Sunny', 'Cloudy', 'Rainy', 'Stormy']`
- Các vector trạng thái thời tiết tương ứng:
  - `sunny = [1,0,0,0]`
  - `cloudy = [0,1,0,0]`
  - `rainy = [0,0,1,0]`
  - `stormy = [0,0,0,1]`

### 6.4.2 Với bộ dữ liệu nhập vào

Bộ dữ liệu gồm:

- Ma trận chuyển trạng thái P được nhập bằng 3 cách sau:
  - Nhập từ bàn phím bằng hàm `input_transition_matrix()`.
  - Nhập trực tiếp từ file csv bằng hàm `get_P('<filename>.csv')`
  - Nhập ma trận số lần chuyển và tính ma trận chuyển bằng hàm `cal_P('<filename>.csv')`
- Vector phân phối đầu được nhập bằng hàm `input_Weather_State_vector()`

### 6.4.3 Các phần dự đoán:

Các phần dự đoán cơ bản giống nhau giữa 2 bộ dữ liệu.

- Dự đoán thời tiết ngày mai dựa trên các trạng thái thời tiết có thể có của hôm nay.
- Dự đoán thời tiết ngày thứ x (kể từ hiện tại) dựa trên các khả năng có thể có của thời tiết có thể có của hôm nay.
- Tìm vector phân phối dừng để dự đoán các xác suất về lâu dài.

## 7 Phụ lục

### 7.1 Định dạng file chứa ma trận chuyển trạng thái

- Dòng đầu tiên và cột đầu tiên chứa các trạng thái.
- Các ô ở dòng i cột j tương ứng là xác suất chuyển từ trạng thái j sang i.

STATISTIC	Sunny	Cloudy	Rainy	Stormy
Sunny	0.4	0.2	0.2	0
Cloudy	0.4	0.4	0.4	0.45
Rainy	0.2	0.3	0.2	0.3
Stormy	0	0.1	0.2	0.25

Bảng 4: Minh hoạ cho file chứa ma trận chuyển trạng thái

## 7.2 Định dạng file chứa ma trận tần suất chuyển trạng thái

- Dòng đầu tiên và cột đầu tiên chứa các trạng thái.
- Các ô ở dòng i cột j tương ứng là tần suất chuyển từ trạng thái j sang i.

COUNT	Sunny	Cloudy	Rainy	Stormy
Sunny	38	36	4	0
Cloudy	34	134	30	10
Rainy	6	26	16	7
Stormy	0	13	5	5

Bảng 5: Minh hoạ cho file chứa ma trận tần suất chuyển

## 7.3 Thư mục lưu các file dữ liệu mẫu:

[Thư mục chứa các file dữ liệu](#): Dữ liệu thời tiết thống kê từ web, file "Weather\_counting.csv" và file "Weather\_statistic.csv" mẫu.

## 8 Tài liệu tham khảo

Các phần tài liệu được cung cấp bởi giảng viên trong quá trình học.