

# Môn học: TOÁN ỨNG DỤNG và THỐNG KÊ

## ĐỒ ÁN THỰC HÀNH CUỐI KỲ

Học kỳ 2, năm học 2023-2024; Lớp: CQ2022

### ĐỀ BÀI

Đồ án thực hành cuối kỳ môn học Toán ứng dụng và thống kê bao gồm 2 câu sau đây:

#### CÂU 1: (5 điểm) Data Fitting.

File “NHANES\_age\_prediction.csv” là bộ dữ liệu được thu thập vào năm 2013-2014 để đánh giá tình trạng sức khỏe và dinh dưỡng của người cao tuổi (từ 65 tuổi trở lên) và người trẻ (dưới 65 tuổi) tại Hoa Kỳ.

Bộ dữ liệu có 2,278 dòng, mỗi dòng là thông tin của 01 người tham gia cuộc khảo sát và có 10 biến (xem mô tả các biến và tải bộ dữ liệu tại link: [https://archive.ics.uci.edu/dataset/887/national+health+and+nutrition+health+survey+2013-2014+\(nhanes\)+age+prediction+subset](https://archive.ics.uci.edu/dataset/887/national+health+and+nutrition+health+survey+2013-2014+(nhanes)+age+prediction+subset)).

Bỏ qua biến “SEQN” (ID của người tham gia khảo sát) và biến “age\_group” (nhóm tuổi mà người tham gia khảo sát thuộc về). Sau đó sử dụng phương pháp hồi quy tuyến tính để xây dựng mô hình dự báo bệnh tiểu đường (biến “DIQ010”).

- Sử dụng toàn bộ 7 biến đầu vào còn lại
- Sử dụng duy nhất 1 biến cho kết quả tốt nhất
- Xây dựng một mô hình của riêng bạn cho kết quả tốt nhất

#### CÂU 2: (5 điểm) Markov chain.

Tòa nhà I của trường Đại học Khoa học Tự nhiên được xây dựng hoàn tất và đưa vào sử dụng từ năm 2000. Giả sử tòa nhà I có tất cả 50 phòng học và cơ sở vật chất (CSVC) của các phòng học này được quản lý bởi Phòng Quản trị thiết bị (PQTTB). PQTTB sẽ khảo sát, thu thập các thông tin về hiện trạng CSVC của các phòng học để đánh giá chất lượng và từ đó lập kế hoạch bảo trì cho phù hợp. Giả sử công tác này được thực hiện hàng năm và hiện trạng CSVC của các phòng học được đánh giá theo 04 loại trạng thái chất lượng sau đây:

- Trạng thái 1: CSVC của phòng học có chất lượng tốt
- Trạng thái 2: CSVC của phòng học có chất lượng khá
- Trạng thái 3: CSVC của phòng học có chất lượng trung bình
- Trạng thái 4: CSVC của phòng học có chất lượng xấu

Giả sử PQTTB đã đánh giá xếp loại chất lượng phòng học trong 2 năm 2020 và 2021 và tổng hợp kết quả như sau:

Trạng thái	Mô tả trạng thái	Số phòng học	
		Năm 2020	Năm 2021
<b>1</b>	CSVC tốt	34	26
<b>2</b>	CSVC khá	12	17
<b>3</b>	CSVC trung bình	3	5
<b>4</b>	CSVC xấu	1	2
Tổng		<b>50</b>	<b>50</b>

Biết rằng trong thời hạn 1 năm, sự xuống cấp về CSVC (thay đổi trạng thái) của các phòng học diễn ra không quá 1 cấp trong điều kiện PQTTB không thực hiện bất kỳ công tác bảo trì, sửa chữa nào. Khi đó, quá trình xuống cấp về CSVC của các phòng học thuộc tòa nhà I là một quá trình ngẫu nhiên có tính Markov. Giả sử năm 2021 là năm bắt đầu của xích Markov.

- Từ bảng số liệu trên, xây dựng ma trận chuyển trạng thái P của xích Markov
- Với các điều kiện của bài toán đã cho, xác định vector phân phối xác suất ban đầu  $\pi_0$
- Hãy dự báo sự xuống cấp về CSVC của các phòng học tòa nhà I trong vòng 15 năm tới, kể từ năm 2021. Kết quả dự báo trình bày trong bảng sau:

Trạng thái		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
Số phòng học trong năm	2021				
	2022				
	2023				
	...				
	2033				
	2034				
	2035				

- Xác định thời điểm (năm) phải phá hủy để xây lại/ trùng tu tòa nhà I (tất cả phòng học đều ở trạng thái CSVC xấu)
- Xác định tuổi thọ của tòa nhà I

## YÊU CẦU CHUNG

- Sinh viên thực hiện bài làm cá nhân (không làm theo nhóm).

- Sinh viên không được dùng thư viện và các hàm có sẵn của các thư viện (chỉ được dùng thư viện pandas cho việc nạp dữ liệu, thư viện matplotlib hoặc seaborn để vẽ hình, thư viện math để thực hiện các tính toán xấp xỉ để khắc phục sai số của máy tính).

## QUY ĐỊNH BÀI NỘP

- Thực hiện mỗi câu trong 01 tập tin Jupyter Notebook (.ipynb) theo quy định sau:
  - Tên tập tin đặt theo cấu trúc: **MSSV\_cau1.ipynb**, **MSSV\_cau2.ipynb**
  - Mở đầu tập tin là đoạn văn bản giới thiệu thông tin cá nhân: Lớp, MSSV, Họ và tên SV
  - Mỗi tập tin mã nguồn chứa toàn bộ bài làm và toàn bộ ví dụ áp dụng theo yêu cầu, Ghi chú ý nghĩa (chức năng) của các hàm tự viết (tại nơi định nghĩa hàm)
- Sinh viên đặt toàn bộ các tập tin bài làm trong 01 tập tin nén **MSSV.zip** hoặc **MSSV.rar**, và nộp tập tin nén này trên website môn học (Moodle) theo đúng thời hạn quy định.

## QUY ĐỊNH CHẤM BÀI

Những trường hợp sau đây sẽ bị 0 điểm toàn bộ bài thi:

- Nộp bài sai quy định
- Thực thi mã nguồn báo lỗi
- **Sao chép bài làm của nhau**

**HẾT**