**Trường đại học Khoa học Tự nhiên**

**Khoa Công nghệ Thông tin**



**BÁO CÁO ĐỒ ÁN MÔ HÌNH MARKOV ẨN (Phần 2: Cài đặt)**

**Môn**: Toán ứng dụng & thống kê

Nhóm sinh viên thực hiện:

20120018 – Trần Kiều Minh Lâm

20120067 – Nguyễn Phạm Bá Duy

20120316 – Nguyễn Hoàng Lâm

Giảng viên hướng dẫn: Nguyễn Đình Thúc, Nguyễn Văn Quang Huy, Võ Nam Thục Đoan

**Năm học:** 2021 – 2022

**1. Cài đặt các thuật toán**

Trình bày trong tập tin algo.py

**1.1 Thuật toán tiến trước (Forward Algorithm)**

**1.2 Thuật toán Viterbi**

**1.3 Thuật toán Baum-Welch**

**2. Giải bài toán**

**a) Mô hình hóa bài toán**

* Không gian mẫu trạng thái ẩn trong mô hình: 2 viên xúc xắc (1 viên bình thường và 1 viên lỗi, gọi lần lượt là A và B).
* Không gian mẫu trạng thái quan sát được: Các mặt số từ 1 – 6.
* Ma trận xác suất ban đầu của trạng thái ẩn (giả sử việc lựa chọn 2 con xúc xắc ban đầu là như nhau).



* Ma trận các xác suất chuyển trạng thái ẩn (transitions):



* Ma trận các xác suất đầu ra trạng thái quan sát được (emissions):



* Chuỗi trạng thái ẩn và chuỗi trạng thái quan sát sẽ được sinh ra từ câu b.

**b) Sinh chuỗi, sử dụng các tham số đề bài cho.**

Trình bày trong tập tin part\_2.py

**c) Dùng Viterbi dự đoán xúc xắc dùng cho mỗi lần tung. Tính độ chính xác.**

Độ chính xác tính bằng số lần kết quả dự đoán và kết quả thực tế giống nhau chia cho 100. Lặp lại các bước trên 10 lần.

**Kết quả:**

Ảnh có chứa văn bản, thiết bị

Mô tả được tạo tự động

Ta thấy độ chính xác của giải thuật Viterbi trong 10 lần đo thấp nhất bằng 0.53 và cao nhất bằng 0.72. Trung bình của 10 lần đo xấp xỉ 0.657. Từ đó có thể nhận xét kết quả dự đoán nằm ở mức chính xác khá cao. (65,7%)

**d) Ước lượng tham số của mô hình.**

Tính A, B dùng Baum-Welch

**Kết quả:**

Ảnh có chứa văn bản

Mô tả được tạo tự động

Thuật toán Baum\_welch cho ra kết quả như trên. Các xác suất chuyển trạng thái ẩn từ A sang [A, B] có sự khác biệt hơi lớn ( [0.8, 0.2] và [0.71, 0.29]). Ngoài ra xác xuất chuyển trạng thái từ B thì gần như nhau ( [0.3, 0.7] và [0.29, 0.71]).

Về ma trận các xác suất đầu ra trạng thái quan sát được, các xác suất của xúc xắc bình thường cũng không bằng nhau như tham số ban đầu mà có sự khác biệt khá lớn. Các xác suất của xúc xắc lỗi có sự khả quan. Ví dụ xác suất ra mặt “6” là 0.53 gần bằng với xác suất ban đầu là 0.5. Các mặt còn lại có xác suất không bằng nhau, từ xấp xỉ 0.0006 cho tới 0.2517

**3. Các tài liệu tham khảo**

* <http://www.adeveloperdiary.com/data-science/machine-learning/forward-and-backward-algorithm-in-hidden-markov-model/>
* <http://www.adeveloperdiary.com/data-science/machine-learning/implement-viterbi-algorithm-in-hidden-markov-model-using-python-and-r/>
* <http://www.adeveloperdiary.com/data-science/machine-learning/derivation-and-implementation-of-baum-welch-algorithm-for-hidden-markov-model/>
* <https://github.com/RRisto/learning/blob/master/markov_chain_learn/baum_welch.ipynb>