Nguyễn Năng Sơn – 16022480

Trình Đức Duy – 17021225

**Sử dụng mô hình HMM để phân loại 5 từ Tiếng Việt**

1. **Thiết kế mô hình**

Mô hình được thiết kế để phân loại 5 từ tiếng Việt là: “đi”, “ra”, “và”, “người”, “y tế”.

* Dữ Liệu

Bộ dữ liệu gồm file âm thanh của 5 từ: “đi”, “ra”, “và”, “người”, “y tế” được tách từ các file ghi âm các câu trong báo cáo 1.

* Trích xuất đặc trưng

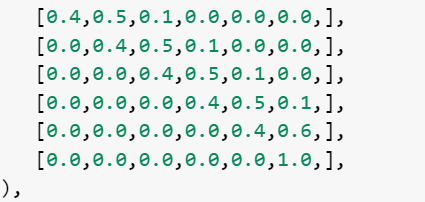
Từ dạng sóng của mỗi file âm thanh, ta chia thành các frame và biến đổi thành các vector mfcc với 12 thành phần.

* Xây dựng mô hình

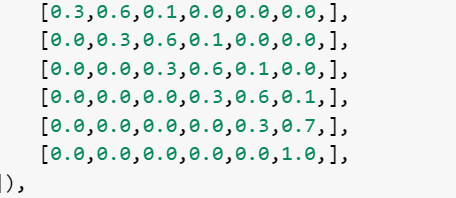
Nhóm sử dụng phương pháp Multinomial HMM, sử dụng gói hmmlearn để triển khai mô hình.

Các mô hình hmm sử dụng kĩ thuật subdividing phones(3 trạng thái cho mỗi âm vị) và 1 skip connection. Phân tích âm vị của 5 từ: mô hình hmm cho từ “đi” có 6 thành phần ứng với 2 âm vị, “ra” có 6 thành phần ứng với 2 âm vị, “và” có 6 thành phần ứng với 2 âm vị, “người” có 12 thành phần ứng với 4 âm vị, “y tế” có 9 thành phần ứng với 3 âm vị.

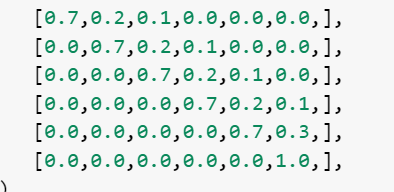
“đi”: Có tổng 2 âm vị nên dùng mô hình có 6 trạng thái



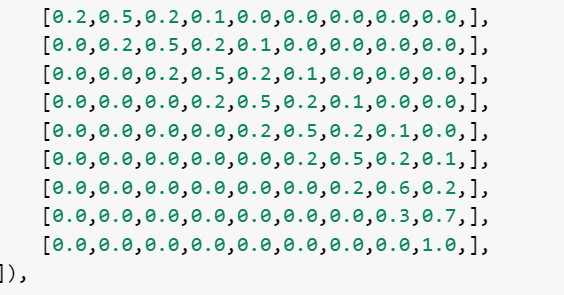
“ra” : Có tổng 2 âm vị nên gồm 6 trạng thái



“và”: Có tổng 2 âm vị nên gồm 6 trạng thái



“ y tế “: Gồm 3 âm vị không tính dấu cách nên có 9 trạng thái



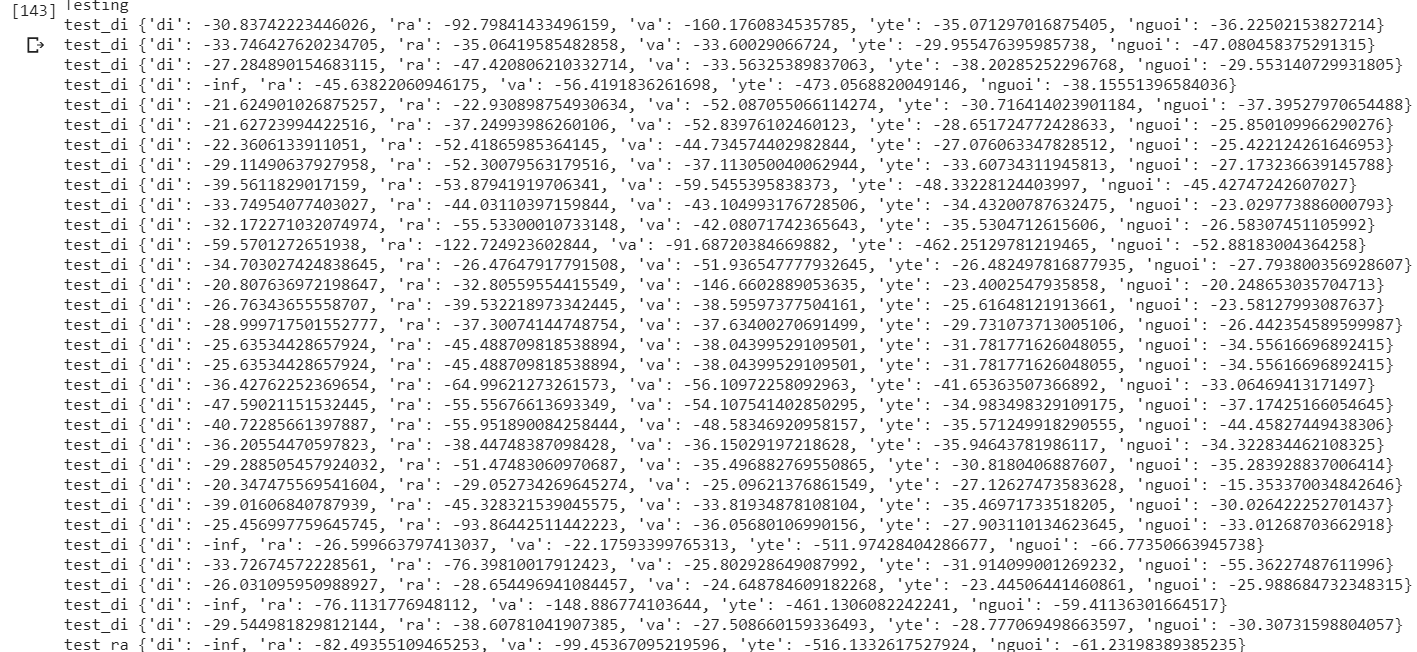
“Người” : Gồm 4 âm vị nên có 12 trạng thái



Mô hình HMM được sử dụng cho mô hình acoustic là left-to-right HMM.

1. **Kiểm tra**

Tập dữ liệu huấn luyện cho mỗi từ có 70 từ, tập test cho mỗi từ sẽ có 30 từ. Các file trong tập test không được huấn luyện trong mô hình.



Kết quả

Testing di : 0.3225806451612903

ra : 0.14285714285714285

va : 0.5666666666666667

yte : 0.95

nguoi : 0.8333333333333334