BÁO CÁO : HUẤN LUYỆN MÔ HÌNH HMM CHO NHẬN DIỆN TIẾNG NÓI

Thành viên:

17021238: Hà Minh Hải

16020196: Nguyễn Đức Anh

1. Mô tả bài toán

Xây dựng 05 mô hình HMM để nhận dạng 05 từ tiếng Việt, trong đó có ít nhất một từ ghép.

2. Dataset

a. Hà nội: 100
b. Học: 100
c. Nhà: 100
d. Nhân viên: 100
e. Tôi: 100

3. Model

Từ	Số state
Hà nội	10
Nhà	6
Học	9
Nhân viên	12
Tôi	9

3.1 Huấn luyện mô hình bằng GMMHMM

Nhóm sử dụng GMMHMM trong gói hmmlearn. Nhóm xây dựng một mô hình hmm từ trái sang phải với các parameter chung:

- n_mix = 5 : thể hiện 5 miền giọng khác nhau trong tập dữ liệu
- n_iter = 1000. Số lần lặp tối đa
- verbose = true.
- params = ‘mctw’. Cho phép huấn luyện m: means; c: covars; t: transmat; w: GMM mixing weights.
- init_params = ‘mct’. Cho phép mô hình tự khởi tạo m, c và t.

3.2 Tôi

- **startprob_**

```
[1.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0]
```

- **transmit_**

```
 [0.7,0.3,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0],
 [0.0,0.7,0.3,0.0,0.0,0.0,0.0,0.0,0.0,0.0],
 [0.0,0.0,0.7,0.3,0.0,0.0,0.0,0.0,0.0,0.0],
 [0.0,0.0,0.0,0.7,0.3,0.0,0.0,0.0,0.0,0.0],
 [0.0,0.0,0.0,0.0,0.7,0.3,0.0,0.0,0.0,0.0],
 [0.0,0.0,0.0,0.0,0.0,0.7,0.3,0.0,0.0,0.0],
 [0.0,0.0,0.0,0.0,0.0,0.0,0.7,0.3,0.0,0.0],
 [0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.7,0.3,0.0],
 [0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.7,0.3],
 [0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,1.0],
```

3.3 Nhà

- **startprob_**

```
[1.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0]
```

- **transmit_**

```
[0.7,0.3,0.0,0.0,0.0,0.0],
[0.0,0.7,0.3,0.0,0.0,0.0],
[0.0,0.0,0.7,0.3,0.0,0.0],
[0.0,0.0,0.0,0.7,0.3,0.0],
[0.0,0.0,0.0,0.0,0.7,0.3],
[0.0,0.0,0.0,0.0,0.0,1.0],
```

3.4 Học

- **startprob_**

```
[1.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0]
```

- **transmit_**

```
 [0.7,0.3,0.0,0.0,0.0,0.0,0.0,0.0,0.0],
 [0.0,0.7,0.3,0.0,0.0,0.0,0.0,0.0,0.0],
 [0.0,0.0,0.7,0.3,0.0,0.0,0.0,0.0,0.0],
 [0.0,0.0,0.0,0.7,0.3,0.0,0.0,0.0,0.0],
 [0.0,0.0,0.0,0.0,0.7,0.3,0.0,0.0,0.0],
 [0.0,0.0,0.0,0.0,0.0,0.7,0.3,0.0,0.0],
 [0.0,0.0,0.0,0.0,0.0,0.0,0.7,0.3,0.0],
 [0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.7,0.3],
 [0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,1.0],
```

3.5 Nhân viên

- **startprob_**

```
[1.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0]
```

- **transmit_**

```
[0.7,0.3,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0],
[0.0,0.7,0.3,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0],
[0.0,0.0,0.7,0.3,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0],
[0.0,0.0,0.0,0.7,0.3,0.0,0.0,0.0,0.0,0.0,0.0,0.0],
[0.0,0.0,0.0,0.0,0.7,0.3,0.0,0.0,0.0,0.0,0.0,0.0],
[0.0,0.0,0.0,0.0,0.0,0.7,0.3,0.0,0.0,0.0,0.0,0.0],
[0.0,0.0,0.0,0.0,0.0,0.0,0.7,0.3,0.0,0.0,0.0,0.0],
[0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.7,0.3,0.0,0.0,0.0],
[0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.7,0.3,0.0,0.0],
[0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.7,0.3,0.0],
[0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.7,0.3],
[0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,1.0],
```

3.6 Hà nội

- **startprob_**

```
[1.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0]
```

- **transmit_**

```
[0.7,0.3,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0],
[0.0,0.7,0.3,0.0,0.0,0.0,0.0,0.0,0.0,0.0],
[0.0,0.0,0.7,0.3,0.0,0.0,0.0,0.0,0.0,0.0],
[0.0,0.0,0.0,0.7,0.3,0.0,0.0,0.0,0.0,0.0],
[0.0,0.0,0.0,0.0,0.7,0.3,0.0,0.0,0.0,0.0],
[0.0,0.0,0.0,0.0,0.0,0.7,0.3,0.0,0.0,0.0],
[0.0,0.0,0.0,0.0,0.0,0.0,0.7,0.3,0.0,0.0],
[0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.7,0.3,0.0],
[0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.7,0.3],
[0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,1.0],
```

4. Kết quả

a. Tập thử nghiệm cắt

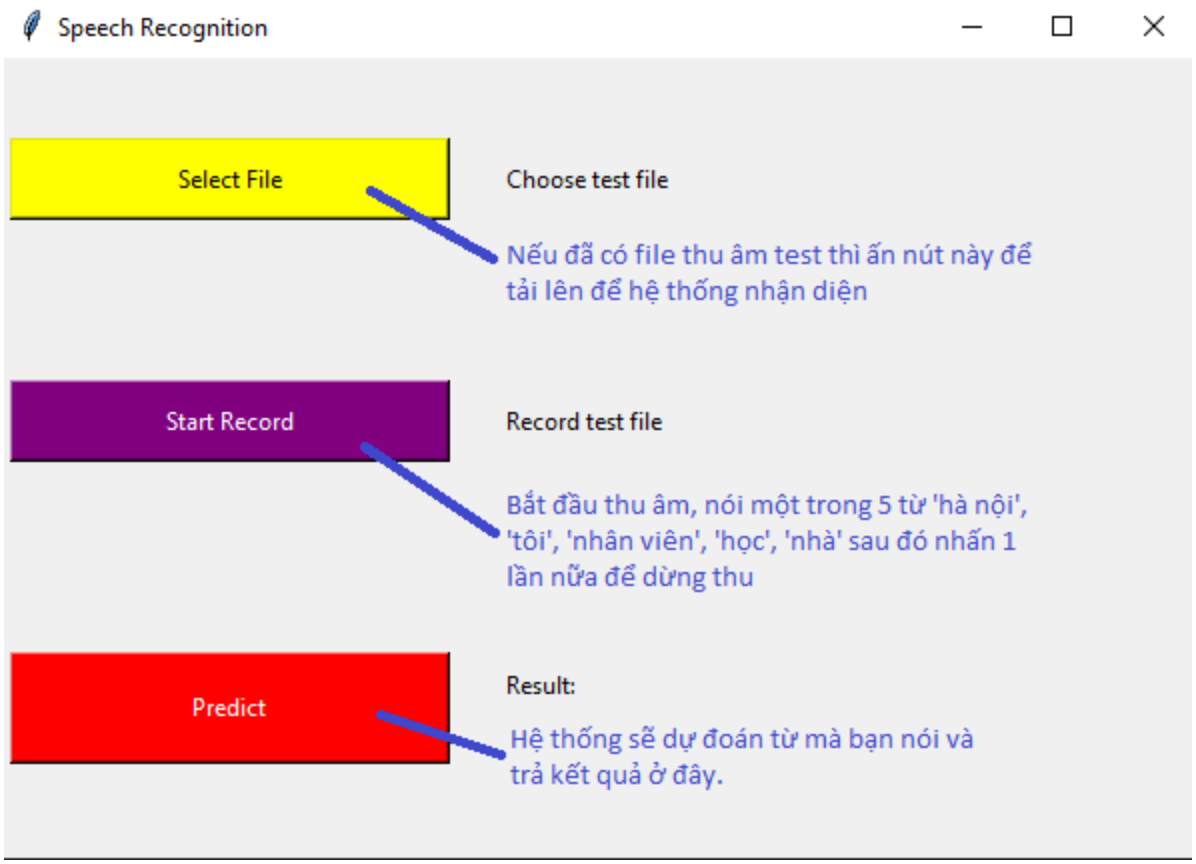
Từ	Dự đoán đúng	Độ chính xác
Tôi	**20/21**	**0.95**
Nhà	**20/20**	**1.0**
Học	**21/21**	**1.0**
Nhân viên	**21/21**	**1.0**
Hà nội	**20/20**	**1.0**

b. Tập thử nghiệm thu trực tiếp

Từ	Dự đoán đúng	Độ chính xác
Tôi	**39/40**	**0.98**
Nhà	**38/40**	**0.95**
Học	**38/40**	**0.95**
Nhân viên	**39/40**	**0.98**
Hà nội	**39/40**	**0.98**

5. Hướng dẫn sử dụng app nhận diện tiếng nói

- Sau khi mở file app.py trong thư mục ta sẽ có cửa sổ hiện lên như sau:


Speech Recognition
Select File
Choose test file
Nếu đã có file thu âm test thì ấn nút này để tải lên để hệ thống nhận diện
Start Record
Record test file
Bắt đầu thu âm, nói một trong 5 từ 'hà nội', 'tôi', 'nhân viên', 'học', 'nhà' sau đó nhấn 1 lần nữa để dừng thu
Predict
Result:
Hệ thống sẽ dự đoán từ mà bạn nói và trả kết quả ở đây.