**BÁO CÁO BÀI TẬP 2: XÂY DỰNG MÔ HÌNH HMM CHO NHẬN DIỆN TIẾNG NÓI**

Thành viên nhóm: Khổng Thị Mai Loan - 17021288

Dương Hải Minh - 17021295

**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. Dữ liệu**

Dữ liệu đầu vào của mô hình HMM là các file âm thanh (file .wav) chứa các từ đơn hoặc từ ghép được cắt ra từ các file ghi âm câu ở bài tập 1.

Ở bài tập này, nhóm sử dụng 05 từ là: “âm tính”, “người”, “tôi”, “trong”, “không”. Đầu tiên, mỗi từ được cắt 100 file làm dữ liệu huấn luyện và thử nghiệm. Hai tập dữ liệu này được chia ngẫu nhiên theo tỷ lệ 80 - 20. Sau khi huấn luyện mô hình xong, một tập dữ liệu thử nghiệm mới được các thành viên nhóm thu âm trực tiếp cũng được sử dụng để đánh giá độ chính xác của mô hình. Số lượng file cụ thể cho tập huấn luyện, tập thử nghiệm và tập thử nghiệm khi thu trực tiếp của mỗi từ như sau:

|  |  |  |  |
| --- | --- | --- | --- |
| Từ | Tập dữ liệu | Tập huấn luyện | Tập kiểm tra |
| Âm tính | 57 | 80% | 20% |
| Người | 710 | 80% | 20% |
| Tôi | 101 | 80% | 20% |
| Trong | 99 | 80% | 20% |
| Không | 100 | 80% | 20% |

**3. Xây dựng mô hình HMM**

Trích xuất đặc trưng MFCC

Mỗi file dữ liệu âm thanh được xử lý để lấy 12 đặc trưng mfcc với win\_length = 25ms và hop\_length = 10ms. Sau đó các đặc trưng mfcc này được chuẩn hóa bằng cách trừ đi giá trị mean của chúng. Sau đó lấy giá trị mfcc sau trừ đi trước để được delta1. Lấy giá trị delta1 sau trừ đi trước để có được delta2. Nối các đặc trưng trên lại, ta có:

12 mfcc + 12 delta1 + 12 delta2 (delta của delta) = 36 đặc trưng. Từ đây, ta nhận được một ma trận X có cỡ T\*36 (với T là số frame) để đưa vào huấn

luyện mô hình hmm.

Xây dựng mô hình bằng GMMHMM

Để xây dựng mô hình hmm, nhóm sử dụng GMMHMM trong gói hmmlearn. Với mỗi từ,

nhóm xây dựng một mô hình hmm từ trái sang phải với các parameter chung:

- n\_mix = 4. Thể hiện 4 miền giọng nói khác nhau trong dữ liệu

- random\_state = 42.

- 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.

**4. Kết quả thử nghiệm**

Kết quả thử nghiệm trên bộ dữ liệu cắt từ ở bài tập 1 và dữ liệu ghi trực tiếp như sau:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Từ | Tập thử nghiệm cắt | | Tập thử nghiệm thu trực tiếp | |
|  | Correct predict | Accuracy | Correct predict | Accuracy |
| Âm tính | 12/12 | 1.0 | 34/40 | 0.85 |
| Người | 129/133 | 0.96 | 40/40 | 1.00 |
| Tôi | 19/21 | 0.90 | 27/40 | 0.675 |
| Có | 19/19 | 1.0 | 40/40 | 1.0 |
| Không | 22/22 | 1.0 | 31/40 | 0.775 |