17021231 – Đỗ Thành Đạt

16020046 – Nguyễn Văn Mạnh

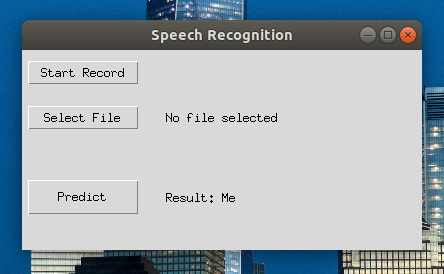
Huấn luyện mô hình HMM

**Mục tiêu:**

Nhận dạng các từ ngữ: Thành Phố, Y Tế, Nhà, Mẹ, Học

**Kiểm thử mô hình:**

Dữ liệu đầu vào được thu âm trực tiếp, sử dụng các model đã được huấn luyện sẵn để nhận dạng từ được thu âm. Các model đưa ra 1 từ dự đoán được giữa 5 từ được huấn luyện, người kiểm thử đưa ra đánh giá dựa trên kết quả nhận dạng

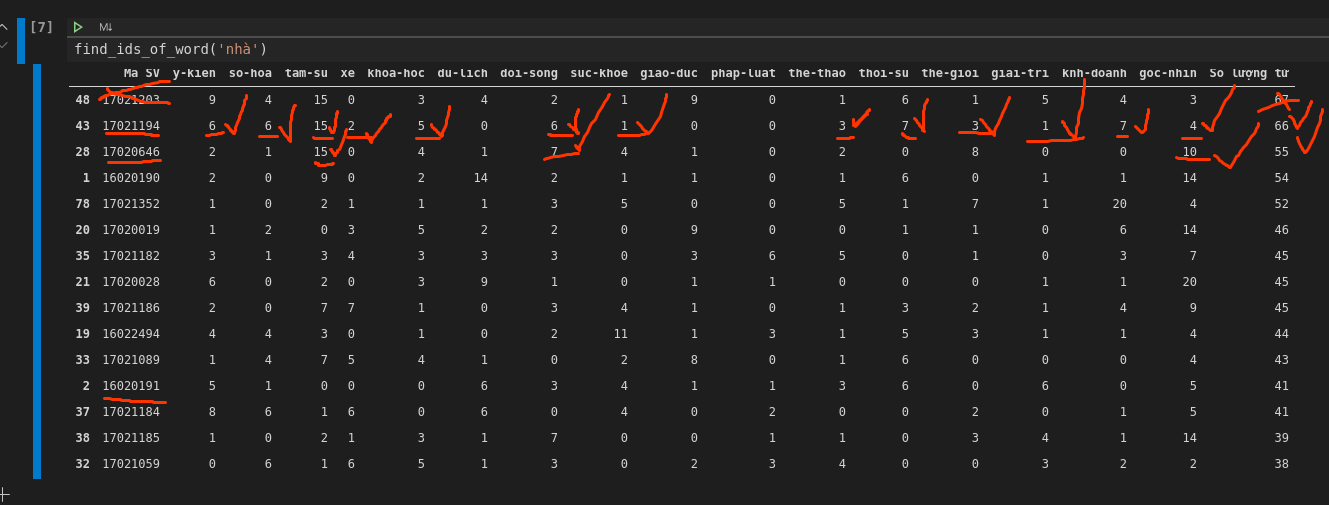
  
Illustration 1: Giao diện ứng dụng nhận dạng

**Dữ liệu huấn luyện:**

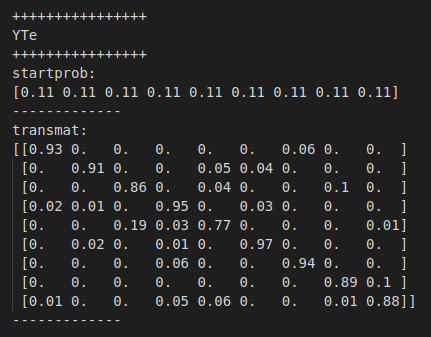
Dữ liệu huấn luyện được cắt từ dữ liệu thu âm các bài báo của cả lớp

Các file chứa từ được chọn được tìm bằng code và dữ liệu transcript cho sẵn. Các file âm thanh được cắt sử dụng Audacity

Với mỗi từ có 100 file âm thanh được tìm và cắt cho việc huấn luyện

  
Illustration 2: Các từ được tìm bằng code và dữ liệu transcript

**Xây dựng mô hình HMM:**

  
Illustration 3: Ví dụ mô hình từ "Y Tế" đã được huấn luyện

Sử dụng hàm GMMHMM của gói thư viện hmmlearn để huấn luyện mô hình HMM. Với mỗi từ được chọn, sử dụng hàm GMMHMM với các thông số:

* n\_mix = 4 (Số state trong GMM)
* n\_iter = 1000 (Số lần lặp tối đa của một mô hình)
* params = “mctw” (Cho phép huấn luyện means, covars, transmat, GMM mixing weight)
* init\_params “mct” (Cho phép mô hình tự khởi tạo means, covars, transmat)
* Ngoài ra còn có các thông số riêng cho từng từ:
  + n\_components: Số state của từng từ đã chọn
  + startprob\_: Khởi tạo Ma trận phân phối ban đầu

Khởi tạo là ma trận có dạng 1xN (N là số states):

[ 1., 0., 0., ……, 0.]

* + transmat\_: Khởi tạo Ma trận chuyển trạng thái giữa các states

Khởi tạo là ma trận có dạng NxN (N là số states), là ma trận đơn vị có đường chéo chính là 1 còn lại là 0:

[[ 1., 0., 0., ……, 0.],

[ 0., 1., 0., ……, 0.],

[ 0., 0., 1., ……, 0.],

……

[ 0., 0., 0., ……, 1.]]

**5 từ đã chọn:**

* **Y tế [** i˧˧ te˧˥ **]**

Tham khảo: *[Y tế (Wikitionary)](https://vi.wiktionary.org/wiki/y_tế" \l "Tiếng_Việt)*

Tham số:

* States: 9
* Kết quả dự đoán:

Trên tập thử nghiệm: 18/20

Thu âm trực tiếp: 16/20

* **Thành phố [** tʰa̤jŋ˨˩ fo˧˥ **]**

Tham khảo: *[Thành phố (Wikitionary)](https://vi.wiktionary.org/wiki/thành_phố" \l "Tiếng_Việt)*

Tham số:

* States: 15
* Kết quả dự đoán:

Trên tập thử nghiệm: 17/20

Thu âm trực tiếp: 16/20

* **Nhà [** ɲa̤ː˨˩ **]**

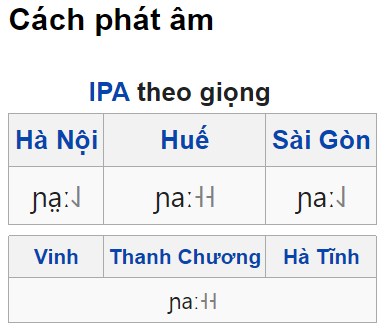
Tham khảo: *[Nhà (Wikitionary)](https://vi.wiktionary.org/wiki/nhà" \l "Tiếng_Việt)*

Tham số:

* States: 6
* Kết quả dự đoán:

Trên tập thử nghiệm: 20/20

Thu âm trực tiếp: 19/20

**Mẹ [** mɛ̰ʔ˨˩ **]**

Tham khảo: *[Mẹ (Wikitionary)](https://vi.wiktionary.org/wiki/mẹ" \l "Tiếng_Việt)*

Tham số:

* States: 6
* Kết quả dự đoán:

Trên tập thử nghiệm: 15/20

Thu âm trực tiếp: 10/20

* **Học [** ha̰ʔwk˨ **]**

Tham khảo: *[Học (Wikitionary)](https://vi.wiktionary.org/wiki/học" \l "Tiếng_Việt)*

Tham số:

* States: 9
* Kết quả dự đoán:

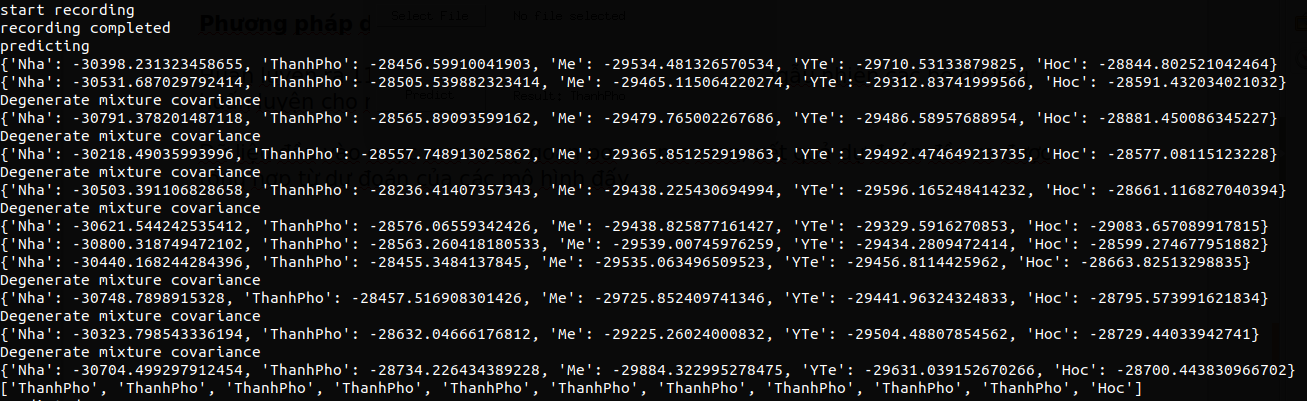
Trên tập thử nghiệm: 18/20

Thu âm trực tiếp: 18/20

**Phương pháp dự đoán:**

Huấn luyện ra 11 mô hình khác nhau, bằng cách chọn ngẫu nhiên các bộ dữ liệu huấn luyện cho mỗi mô hình

Dữ liệu đầu vào được đưa ra dự đoán bởi 11 mô hình, kết quả dự đoán đầu ra được tổng hợp từ dự đoán của các mô hình đấy

  
Illustration 4: Các mô hình đưa ra dự đoán