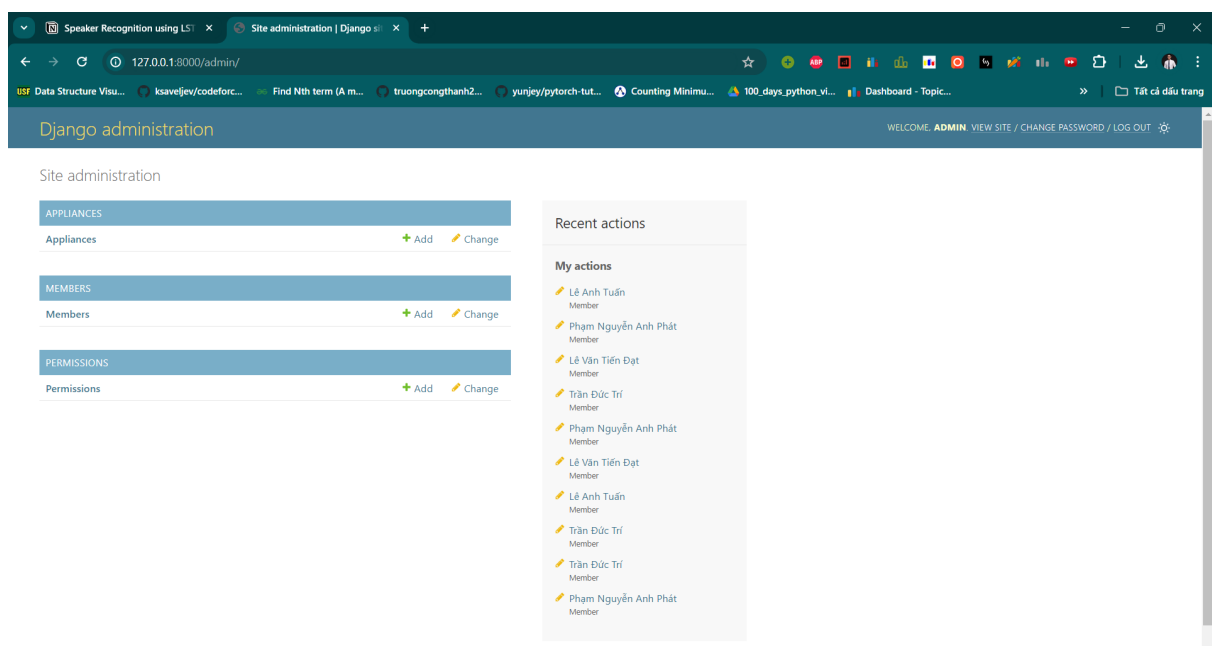


Speaker Recognition using LSTM Implementation on Raspberry Pi 4

I. Đặt vấn đề và phương pháp giải quyết

- Đồ án hiện đảm bảo yêu cầu về AI cũng như phần cứng, tuy nhiên hiện đang thiếu hụt về lĩnh vực IOT
- Xây dựng website BackEnd giải quyết các vấn đề enroll thành viên mới trong gia đình, phân quyền điều khiển thiết bị cũng như linh hoạt quản lý nhà thông minh.



Speaker Recognition using LSTM

Select appliance to change | D

127.0.0.1:8000/admin/appliances/appliance/

Data Structure Visu...

ksaveljev/codeforc...

Find Nth term (A m...

truongcongthanh2...

yunjeey/pytorch-tut...

Counting Minimu...

100_days_python.vi...

Dashboard - Topic...

Tất cả dấu trang

Start typing to filter...

APPLIANCES

Appliances + Add

MEMBERS

Members + Add

PERMISSIONS

Permissions + Add

Select appliance to change

ADD APPLIANCE +

Q

Search

Action: -----

Go

0 of 10 selected

<input type="checkbox"/>	ID	NAME	DESCRIPTION
<input type="checkbox"/>	11	Cảm biến nhiệt độ độ ẩm phòng bếp	-Tên: Cảm biến nhiệt độ độ ẩm phòng bếp -Sơ đồ mạch: + Chấn: + Chấn: -Quyền: + Lê Anh Tuấn + Phạm Nguyễn Anh Phát + Trần Đức Trí + Lê Văn Tiến Đạt
<input type="checkbox"/>	10	Đèn phòng bếp	-Tên: Đèn phòng bếp -Sơ đồ mạch: + Chấn: + Chấn: -Quyền: + Lê Anh Tuấn + Phạm Nguyễn Anh Phát + Trần Đức Trí + Lê Văn Tiến Đạt
<input type="checkbox"/>	9	Đèn phòng ngủ con cái	-Tên: Đèn phòng ngủ con cái -Sơ đồ mạch: + Chấn: + Chấn: -Quyền: + Lê Anh Tuấn + Phạm Nguyễn Anh Phát + Trần Đức Trí + Lê Văn Tiến Đạt
<input type="checkbox"/>	8	Cửa phòng ngủ con cái	-Tên: Cửa phòng ngủ con cái -Sơ đồ mạch: + Chấn: + Chấn: -Quyền: + Lê Anh Tuấn + Phạm Nguyễn Anh Phát + Trần Đức Trí + Lê Văn Tiến Đạt
<input type="checkbox"/>	7	Đèn phòng ngủ ba mẹ	-Tên: Đèn phòng ngủ ba mẹ -Sơ đồ mạch: + Chấn: + Chấn: -Quyền: + Lê Anh Tuấn + Phạm Nguyễn Anh Phát
<input type="checkbox"/>	6	Cửa phòng ngủ ba mẹ	-Tên: Cửa phòng ngủ ba mẹ -Sơ đồ mạch: + Chấn: + Chấn: -Quyền: + Lê Anh Tuấn + Phạm Nguyễn Anh Phát
<input type="checkbox"/>	5	Đèn phòng khách	-Tên: Đèn phòng khách -Sơ đồ mạch: + Chấn: + Chấn: -Quyền: + Lê Anh Tuấn + Phạm Nguyễn Anh Phát + Trần Đức Trí + Lê Văn Tiến Đạt
<input type="checkbox"/>	4	Cửa trượt phòng khách	-Tên: Cửa trượt phòng khách -Sơ đồ mạch: + Chấn: + Chấn: -Quyền: + Lê Anh Tuấn + Phạm Nguyễn Anh Phát + Trần Đức Trí + Lê Văn Tiến Đạt
<input type="checkbox"/>	3	Cửa cuốn garage	-Tên: Cửa cuốn garage -Sơ đồ mạch: + Chấn: + Chấn: -Quyền: + Lê Anh Tuấn + Phạm Nguyễn Anh Phát
<input type="checkbox"/>	2	Đèn garage	-Tên: Đèn trong garage -Sơ đồ mạch: + Chấn: + Chấn: -Quyền: + Lê Anh Tuấn + Phạm Nguyễn Anh Phát + Trần Đức Trí + Lê Văn Tiến Đạt

Filter

By name

All

Cảm biến nhiệt độ độ ẩm phòng bếp

Cửa cuốn garage

Cửa phòng ngủ ba mẹ

Cửa phòng ngủ con cái

Cửa trượt phòng khách

Đèn garage

Đèn phòng bếp

Đèn phòng khách

Đèn phòng ngủ ba mẹ

Đèn phòng ngủ con cái

Speaker Recognition using LSTM

Phạm Nguyễn Anh Phát | Cha

127.0.0.1:8000/admin/members/member/5/change/

Data Structure Visu...

ksaveljev/codeforc...

Find Nth term (A m...

truongcongthanh2...

yunjeey/pytorch-tut...

Counting Minimu...

100_days_python.vi...

Dashboard - Topic...

Tất cả dấu trang

Start typing to filter...

APPLIANCES

Appliances + Add

MEMBERS

Members + Add

PERMISSIONS

Permissions + Add

Change member

HISTORY

Phạm Nguyễn Anh Phát

Name:

Phạm Nguyễn Anh Phát

About:

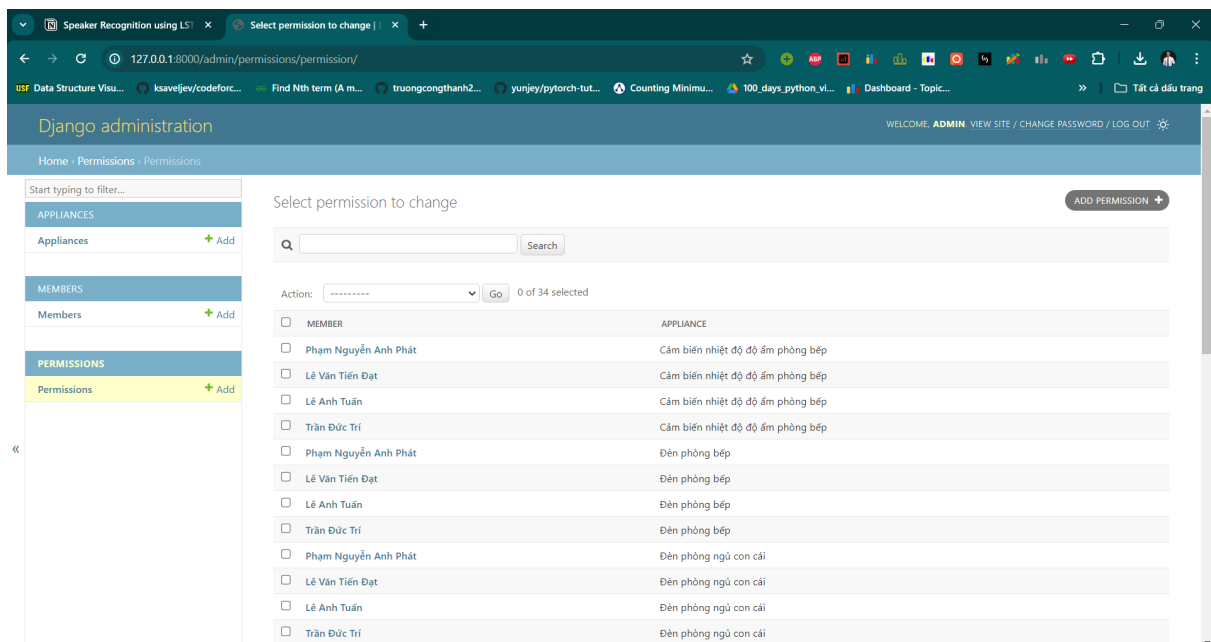
phó bí thư (ba mẹ)

MEMBER FILES

FILE	DELETE?
<div>Phạm Nguyễn Anh Phát - File</div> <div>Currently: audio_raw_data/Phạm Nguyễn Anh Phát/phanh_laptop_1.m4a</div> <div>Change: <input type="button" value="Chọn tệp"/> Không có tệp nào được chọn</div>	<input type="checkbox"/>
<div>Phạm Nguyễn Anh Phát - File</div> <div>Currently: audio_raw_data/Phạm Nguyễn Anh Phát/phanh_laptop_2.m4a</div> <div>Change: <input type="button" value="Chọn tệp"/> Không có tệp nào được chọn</div>	<input type="checkbox"/>
<div>Phạm Nguyễn Anh Phát - File</div> <div>Currently: audio_raw_data/Phạm Nguyễn Anh Phát/phanh_laptop_3.m4a</div> <div>Change: <input type="button" value="Chọn tệp"/> Không có tệp nào được chọn</div>	<input type="checkbox"/>

Speaker Recognition using LSTM Implementation on Raspberry Pi 4

2



II. Triển khai Speaker Recognition với âm thanh được ghi âm trực tiếp từ Laptop và micro

- Thống kê kết quả từ bài Report trước:

Thống kê kết quả tốt nhất của các phương pháp:

	Sử dụng Model thuần túy	Sử dụng KNN	Sử dụng KMean
Data Tổng hợp	83.333	93.137	92.16
Data Tiếng nói để train (20 giây)	97.5	97.5	100.0
Data Tiếng nói điều khiển nhà (3 giây)	74.193	90.323	87.10

Vấn đề mà nhóm gặp phải:

- Khi thu âm trực tiếp trên máy tính, chẳng hạn như khi ghi âm với các câu "Mở cửa phòng ngủ " thì kết quả chạy không chính xác. Điểm chung những câu này là thời lượng ngắn (độ dài trung bình dưới 1-2 giây) cũng như âm thanh để tạo ra base embedding vector cho người nói cũng chỉ kéo dài 5s. (Tần suất xuất hiện của người nói đúng là 2/5-3/5 \Rightarrow không mang nhiều ý nghĩa và ổn định)

Sau khi tham khảo các diễn đàn về MFCC thì tầm quan trọng không phải **độ dài audio** mà là **số mẫu lấy được trên audio** đó. Từ đó bọn em đề xuất hướng giải quyết vấn đề nêu trên:

Hướng giải quyết :

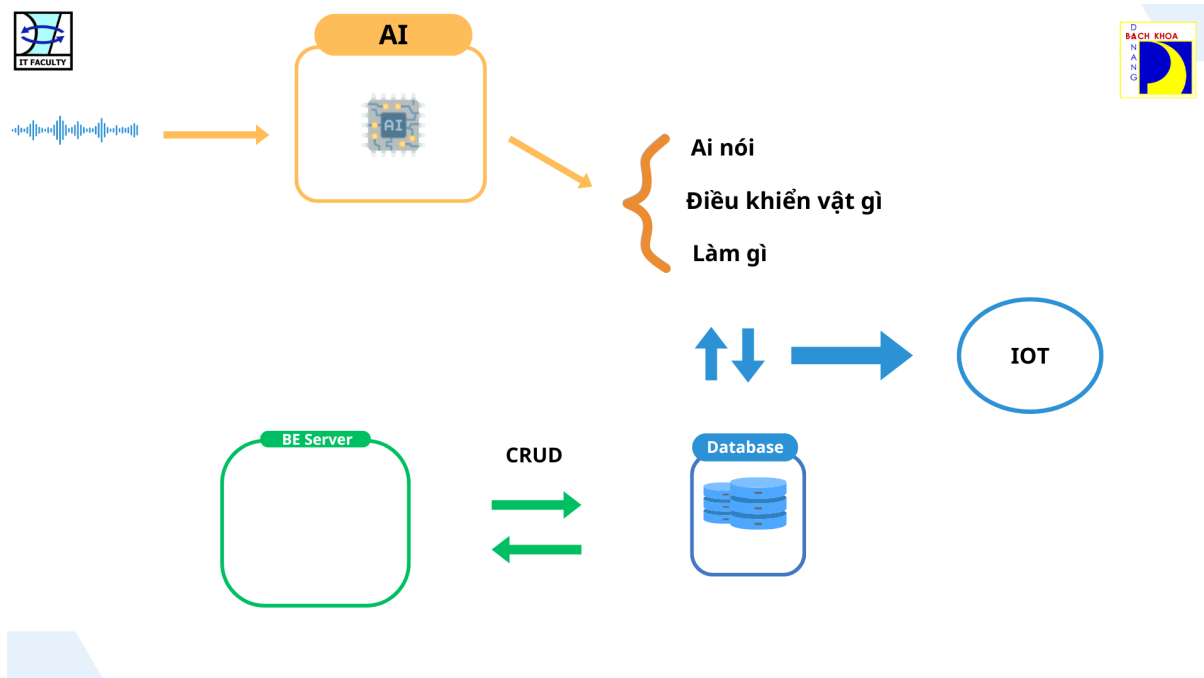
- Từ đoạn audio thu âm được, tạo thành chính đoạn video đó lặp lại nhiều lần, kết quả có cải thiện nhưng chưa cao.
- Thêm vào đó ta cần phải nói rõ và nói nhiều, tức làm sao để số đặc trưng từ mẫu MFCC có được đa dạng và nhiều, tránh các khoảng lặng nhiều để có thể gần giống với đặt trưng từ mẫu gốc nhất có thể.

Kết quả thực nghiệm nhóm thu nhận được:

- Tần suất người nói đúng xuất hiện đa phần từ 4/5 đến 5/5 (K_Neighbours bọn em chọn bằng 5)
- Kết quả có độ chính xác tương đối cao, thực nghiệm cho thấy đạt được kết quả **37/40** lần thu âm trả về người nói chính xác.

III. Triển khai Speaker Recognition với âm thanh được ghi âm trực tiếp từ Raspberry và micro

- **Nhắc lại luồng hoạt động của hệ thống**



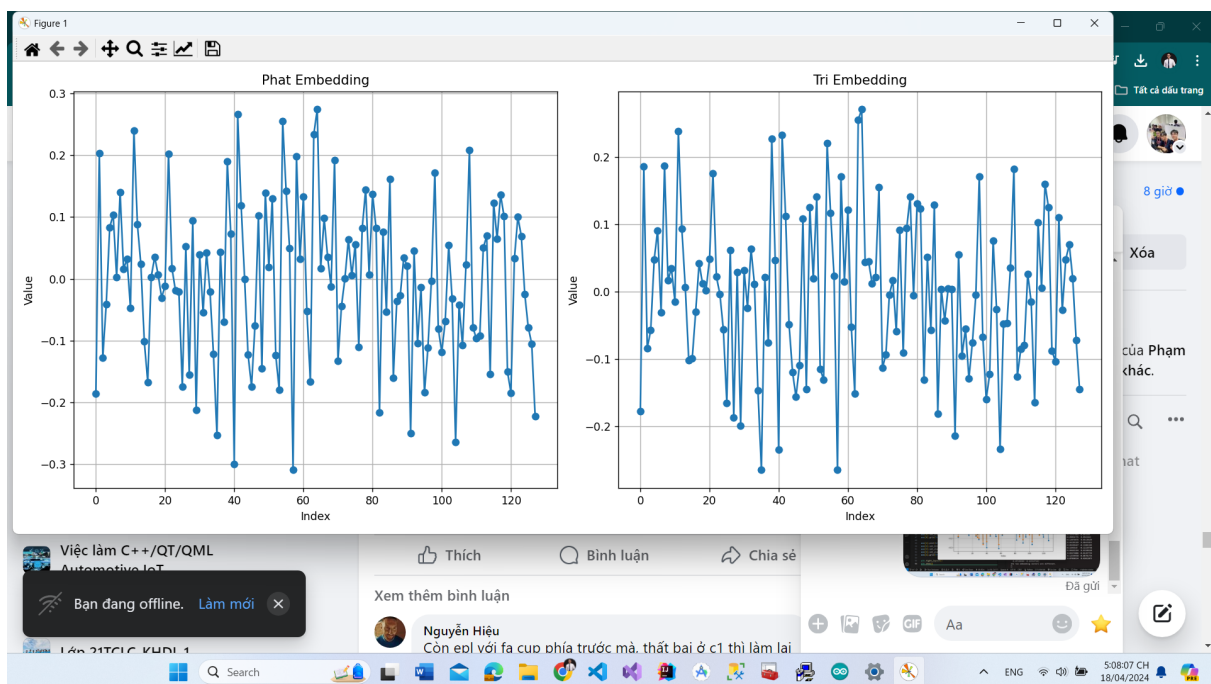
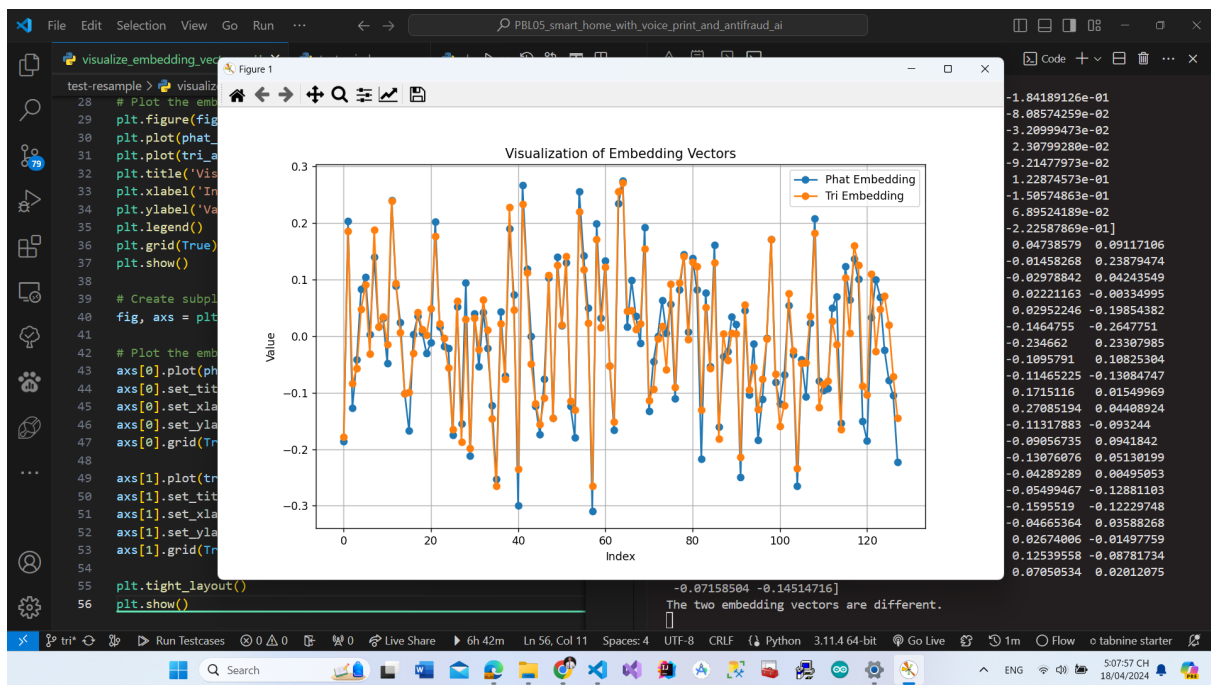
- Nhóm bắt đầu triển khai phần mềm trên Raspberry từ thứ 7 (13/4/2024) và làm việc liên tục nhưng vẫn còn gặp vấn đề nan giải, mong nhận được sự hỗ trợ và tư vấn từ thầy Duy.
- Nhóm triển khai Code AI, thuật toán, dữ liệu y hệt như khi đang chạy local ở laptop, tuy nhiên kết quả không được như mong muốn.

Vấn đề mà nhóm gặp phải:

- Kết quả chương trình trả về luôn là 5/5 người nói nhận diện được là Trần Đức Trí với bất kỳ người thực hiện ghi âm là ai.

Hướng giải quyết:

- Chúng em đã xem xét rất kỹ từng dòng code, kiểm tra lại mặt logic cũng như dữ liệu nhưng vẫn chưa tìm được nguyên nhân cốt lõi của vấn đề trên. Một vài trong số những nguyên nhân mà nhóm đã tìm hiểu như sau:
- Code khi bọn em chạy dưới laptop và khi đưa lên Raspberry, hàm Resampling giữa các thư viện khác nhau thì trả về kết quả khác nhau. Tuy nhiên khi sử dụng chung một hàm thì kết quả vẫn bị sai lệch.



- Phần mềm thu âm: Độ chính xác đạt được 93.33% như đã đề cập ở báo cáo trước khi nhóm xử lý với những file thu âm bằng điện thoại Iphone với chất lượng âm thanh cực kỳ tốt. Và đương nhiên, khi sử dụng base embedding vector là một nguồn âm tốt như Iphone để nhận diện âm thanh thu bằng Micro (cho dù được kết nối với Rasp hay Laptop) thì đều nhận diện sai và ngược lại.
- Tuy nhiên, điều mâu thuẫn nhất mà nhóm vẫn chưa trả lời được đó là chất lượng âm thanh khi thu âm bằng Rasp hay Laptop thì chất lượng ngang

nhau, cũng như các thông số tương tự. Tuy nhiên Lap thì trả về kết quả đúng còn Rasp thì trả về kết quả sai hoàn toàn (Luôn trả về Trần Đức Trí).

Nhóm đã cùng nhau làm việc trong 4 ngày nhưng vấn đề vẫn còn chưa được giải quyết. Nếu xong thì sẽ triển khai phần cứng rất nhanh.

IV. Hướng giải quyết tạm thời

- Tuấn đang sử dụng một model khác **wav2vec** và hiện đạt được những kết quả khả quan.
- Nhóm sẽ tiếp tục nghiên cứu cách tiền xử lý âm thanh đầu vào khi thu âm qua micro truyền vào Raspberry.