**`TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**

**KHOA CÔNG NGHỆ THÔNG TIN**

**BỘ MÔN CÔNG NGHỆ PHẦN MỀM**

**TRẦN NGỌC QUANG – NGUYỄN HOÀNG QUYÊN**

**XÂY DỰNG MÔ HÌNH**

**NHẬN DẠNG ÂM THANH TIẾNG VIỆT**

**KHÓA LUẬN TỐT NGHIỆP CỬ NHÂN CNTT**

**TP.HCM, 2020**

**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**

**KHOA CÔNG NGHỆ THÔNG TIN**

**BỘ MÔN CÔNG NGHỆ PHẦN MỀM**

**TRẦN NGỌC QUANG - 1712706**

**NGUYỄN HOÀNG QUYÊN - 1712712**

**XÂY DỰNG MÔ HÌNH**

**NHẬN DẠNG ÂM THANH TIẾNG VIỆT**

**KHÓA LUẬN TỐT NGHIỆP CỬ NHÂN CNTT**

GIÁO VIÊN HƯỚNG DẪN

**TS. NGÔ HUY BIÊN**

**KHOÁ 2017 - 2021**

**NHẬN XÉT CỦA GIÁO VIÊN HƯỚNG DẪN**

................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................

TpHCM, ngày . . . tháng . . . năm 2021

Giáo viên hướng dẫn

[Kí tên và ghi rõ họ tên]

**NHẬN XÉT CỦA GIÁO VIÊN PHẢN BIỆN**

................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................

Khoá luận đáp ứng yêu cầu của Khoá luận cử nhân CNTT.

TpHCM, ngày . . . tháng . . . năm 2021

Giáo viên phản biện

[Kí tên và ghi rõ họ tên]

Khoa Công Nghệ Thông Tin

Bộ môn Công Nghệ Phần Mềm

**ĐỀ CƯƠNG CHI TIẾT**

|  |  |
| --- | --- |
| **Tên đề tài:** Xây dựng mô hình nhận dạng âm thanh tiếng Việt | |
| **Giáo viên hướng dẫn:** TS. Ngô Huy Biên | |
| **Thời gian thực hiện:** 14/09/2020 đến 29/03/2021 | |
| **Sinh viên thực hiện:** Trần Ngọc Quang – 1712706, Nguyễn Hoàng Quyên - 1712712 | |
| **Loại đề tài:** Nghiên cứu | |
| **Nội dung đề tài:**   1. Trình bày lý do xây dựng mô hình nhận dạng âm thanh tiếng Việt. 2. Trình bày lý thuyết nền tảng và giải pháp để xử lý việc nhận một tập tin âm thanh tiếng Việt và xuất ra nội dung văn bản ở dạng Tiếng Việt. 3. Xây dựng, thu thập dữ liệu, và đào tạo mô hình để nhận một tập tin âm thanh tiếng Việt và xuất ra nội dung ở dạng văn bản 4. Xây dựng dịch vụ web (API) để nhận một tập tin âm thanh tiếng Việt và xuất ra nội dung ở dạng văn bản. 5. Website mẫu việc sử dụng API của mô hình dịch máy từ tiếng Anh sang tiếng Việt đã xây dựng. 6. Cải tiến độ chính xác của mô hình với mục tiêu là 75%. 7. Đề xuất giải pháp cải thiện độ chính xác của mô hình và định hướng mở rộng cho ứng dụng của mô hình. | |
| **Kế hoạch thực hiện:**   |  |  |  | | --- | --- | --- | | **Thời gian thực hiện** | **Công việc thực hiện** | **Người thực hiện** | | 7/9/2020-13/9/2020 | - Nhận đề tài  - Xây dựng bản kế hoạch sơ bộ cho các công việc cần thực hiện | Quang, Quyên | | 14/9/2020-20/9/2020 | - Tìm hiểu và phân tích các yêu cầu về kiến thức nền cho đề tài.  - Khảo sát, dùng thử các hệ thống cung cấp dịch vụ mẫu có sẵn trên thị trường: fpt.ai, vais.vn,...  - Tạo Trello | Quang, Quyên | | 21/9/2020-27/9/2020 | - Thống nhất nội dung chính của ứng dụng demo việc sử dụng API.  - Biên soạn đề cương cho luận văn (dạng slide). | Quang, Quyên | | 28/9/2020-11/10/2020 | - Tìm hiểu lý thuyết nền tảng trong máy học.  - Biên soạn đề cương cho luận văn (dạng word).  - Tìm hiểu lý thuyết nền tảng trong việc nhận dạng giọng nói. | Quang, Quyên | | 14/10/2020-1/11/2020 | - Tìm hiểu về các thư viện Scikit-Learn, Tensorflow, Keras.  - Viết chương 1 luận văn.  - Tìm hiểu các model và kiến trúc, chạy thử các ví dụ để đánh giá.  - Chỉnh sửa chương 1 luận văn. | Quang, Quyên | | 2/11/2020-8/11/2020 | - Chạy thử mô hình nhận dạng âm thanh Tiếng Việt sang văn bản. | Quang, Quyên | | 9/11/2020-22/11/2020 | - Thu thập dữ liệu âm thanh.  - Chỉnh sửa dữ liệu âm thanh. | Quang, Quyên | | 23/11/2020-29/11/2020 | - Tìm hiểu và xây dựng mô hình nhận dạng âm thanh Tiếng Việt. | Quang, Quyên | | 30/11/2020-13/12/2020 | - Tiếp tục xây dựng mô hình nhận dạng âm thanh Tiếng Việt. | Quang, Quyên | | 14/12/2020-20/12/2020 | - Huấn luyện mô hình.  - Viết chương 2 luận văn. | Quang, Quyên | | 21/12/2020-3/1/2020 | - Cải tiến mô hình.  - Chỉnh sửa chương 2 luận văn. | Quang, Quyên | | 4/1/2020-17/1/2020 | - Viết chương 3 luận văn.  - Chỉnh sửa chương 3 luận văn. | Quang, Quyên | | 18/1/2020-31/1/2020 | - Xây dựng và triển khai hệ thống cung cấp dịch vụ web (API).  - Viết chương 4 luận văn. | Quang, Quyên | | 1/2/2020-7/2/2020 | - Xây dựng ứng dụng demo việc sử dụng API trên nền tảng web.  - Chỉnh sửa chương 4 luận văn. | Quang, Quyên | | 8/2/2020-14/2/2020 | - Viết chương 5 luận văn.  - Chỉnh sửa chương 5 luận văn. | Quang, Quyên | | 15/2/2020-21/2/2020 | - Hoàn thành luận văn.  - Chỉnh sửa và cải thiện hiệu năng ứng dụng demo.  - Nâng cấp mô hình hoàn thiện hơn.  - Cải thiện hiệu năng hệ thống cung cấp dịch vụ web (API). | Quang, Quyên | | 22/2/2020-28/2/2020 | - Hoàn chỉnh cuốn luận văn. | Quang, Quyên | | 1/3/2020-7/3/2020 | - Hoàn chỉnh slide trình bày.  - Hoàn chỉnh sản phẩm khoá luận. | Quang, Quyên | | |
| **Xác nhận của giáo viên hướng dẫn** | **Ngày . . . tháng . . . năm 2020**  **Sinh viên thực hiện** |

MỤC LỤC

[CHƯƠNG 1: GIỚI THIỆU LUẬN VĂN 2](#_Toc53569834)

[1.1 GIỚI THIỆU ĐỀ TÀI 2](#_Toc53569835)

[1.1.1 Phục vụ nhu cầu trong cuộc sống hằng ngày 3](#_Toc53569836)

[1.1.2 Áp dụng vào nhiều lĩnh vực trong cuộc sống 3](#_Toc53569837)

[1.1.3 Tiềm năng kinh tế cao 4](#_Toc53569838)

[1.1.4 Phục vụ nhu cầu sử dụng công nghệ cho đa dạng người trong xã hội 5](#_Toc53569839)

[1.2 LÝ DO LỰA CHỌN ĐỀ TÀI 6](#_Toc53569840)

[1.3 HƯỚNG PHÁT TRIỂN CỦA LUẬN VĂN 6](#_Toc53569841)

[1.4 MỤC TIÊU CỦA LUẬN VĂN 7](#_Toc53569842)

[1.5 PHẠM VI ĐỀ TÀI 8](#_Toc53569843)

[TÀI LIỆU THAM KHẢO 10](#_Toc53569844)

# **CHƯƠNG 1: GIỚI THIỆU LUẬN VĂN**

# **GIỚI THIỆU ĐỀ TÀI**

Cuộc sống ngày càng hiện đại, nhu cầu áp dụng những công nghệ tiên tiến để phục vụ các công việc trong cuộc sống của mỗi người càng cao. Hơn một thập kỷ gần đây, với sự phát triển mạnh mẽ của công nghệ thông tin, công nghệ xử lý ngôn ngữ tự nhiên như mã hóa, nhận dạng tiếng nói, giả lập giọng nói…đã trở thành xu hướng nghiên cứu mới được nhiều nhà khoa học quan tâm ở các lĩnh vực khác nhau như tin học, toán học, điều khiển, điện tử,…

Với xu hướng mới đó, vấn đề giao tiếp giữa con người và máy tính đang đòi hỏi những sự cải tiến vượt bậc. Công nghệ đã làm tăng lên tốc độ xử lý dữ liệu của các thiết bị hiện đại trong cuộc sống hàng ngày của chúng ta. Cùng với sự tăng lên của tốc độ xử lý dữ liệu, thời gian mà máy móc đợi con người nhập liệu cũng tăng lên. Nhìn chung, hầu hết các phương pháp nhập liệu truyền thống, có thể kể đến như nhập liệu bằng tay, nhập liệu thông qua các thiết bị điều khiển, ... có tốc độ và thời gian thao tác nhập dữ liệu tương đối chậm. Để giải quyết vấn đề này, giọng nói là giải pháp hợp lý nhất tính đến thời điểm hiện tại. Với sự trợ giúp của công nghệ, người dùng có thể dễ dàng điều khiển thiết bị và nhập tài liệu bằng giọng nói. Đó là chưa kể đến lợi ích to lớn mà những phần mềm dựa trên nhận dạng giọng nói có thể mang lại cho những người khiếm thị. Họ có thể tận hưởng những tiến bộ công nghệ tương tự như những gì mà một người bình thường có thể làm, không còn khoảng cách xuất hiện do những khiếm khuyết về giác quan. Nhận dạng giọng nói cho phép tạo tài liệu nhanh hơn vì phần mềm nói chung tạo ra các từ thông qua giọng nói thường nhanh hơn nhiều so với một người có thể nhập. Có thể nói, giọng nói đang ngày càng trở nên "quyền lực" hơn khi có thể điều khiển mọi thiết bị công nghệ hiện đại.

Và thực tế, nhận dạng giọng nói tự động đang dần bùng nổ, đem lại nhiều lợi ích và thuận lợi hơn cho con người. Nhận thấy được tiềm năng phát triển to lớn như vậy, hàng loạt các dịch vụ nhận dạng âm thanh tiếng nói cũng như là ứng dụng sử dụng dịch vụ này đã ra đời. Từ các tổ chức đến cá nhân, công nghệ này được sử dụng rộng rãi vì những lợi thế khác nhau mà nó mang lại.

Sau đây nhóm sinh viên sẽ nêu qua một số lĩnh vực mà công nghệ này mang lại.

## **1.1.1 Phục vụ nhu cầu trong cuộc sống hằng ngày**

Trong cuộc sống hằng ngày, việc sử dụng công nghệ nhận dạng giọng nói giúp giảm tiết kiệm nhiều thời gian, sức lực cho con người. Có thể kể đến như : điều khiển các thiết bị trong nhà thông qua giọng nói; tìm kiếm thông tin, tìm kiếm kênh trên tivi, máy tinh,... mà không cần nhập tay; nhắn tin với bạn bè khi đang bận một công việc khác; ghi chú lại các thông tin trong cuộc hội thảo, cuộc nói chuyện mà không cần tốn thời gian bằng việc ghi lại bằng tay,... Các trợ lý ảo trên điện thoại hoặc các mạch nhận dạng giọng nói trên các thiết bị đang thực hiện tốt các nhiệm vụ trên giúp con người có trải nghiệm tốt nhất trong cuộc sống thường nhật.

## **1.1.2 Áp dụng vào nhiều lĩnh vực trong cuộc sống**

Công nghệ nhận dạng giọng nói và việc sử dụng trợ lý ảo đã nhanh chóng được áp dụng vào nhiều lĩnh vực khác nhau trong cuộc sống, chẳng hạn như văn phòng, kinh doanh, y tế, vạn vật kết nối, giáo dục, ...

* **Trong lĩnh vực văn phòng**

Công nghệ nhận dạng giọng nói ở lĩnh vực văn phòng đã phát triển mạnh, có vai trò kết hợp các nhiệm vụ đơn giản để tăng hiệu quả, giảm các công việc truyền thống do con người thực hiện để tăng năng suất.

Các công việc mà trợ lý ảo có thể thực hiện được như: tìm kiếm tài liệu, văn bản, báo cáo,... trên thiết bị; nhận dạng thông tin muốn đưa vào máy tính; in, xuất tài liệu; thực hiện các công việc trong cuộc hội thảo; tạo các biểu đồ, bảng báo cáo bằng dữ liệu,... Tất cả các yêu cầu trên đều có thể giao tiếp với trợ lý ảo bằng giọng nói.

* **Trong lĩnh vực kinh doanh**

Nhận dạng giọng nói giúp tăng cường nhu cầu tìm kiếm bằng giọng nói, từ đó mở ra một khía cạnh mới mà nhà tiếp thị có thể tiếp cận người tiêu dùng. Dữ liệu giọng nói của người dùng có thể giải thích độ tuổi, tầng lớp và các thông tin liên quan đến sinh trắc nhân khẩu như văn hóa của họ. Các nhà tiếp thị có thể tối ưu hóa những thông tin này để đón đầu các xu hướng tiêu dùng.

* **Trong lĩnh vực y tế**

Đặc điểm của công nghệ nhận dạng giọng nói là giảm thiểu thời gian giao tiếp với thiết bị công nghệ. Do đó, việc áp dụng công nghệ này vào các bệnh viện sẽ tiết kiệm thời gian quý giá bởi việc truy cập thông tin rảnh tay, mang lại ích lợi vô cùng tích cực đến sự an toàn của bệnh nhân, nâng cao hiệu quả y tế. Một số thuận lợi cụ thể của nhận dạng giọng nói trong lĩnh vực ý tế như: tìm các hồ sơ y tế nhanh chóng; truy vấn thông tin về quản lý như số giường trống, số bệnh nhân hiện có, ... cải thiện quy trình làm việc; rút gọn thời gian nhập liệu; trong tương lai có thể có sự tương tác giữa bệnh nhân và trợ lý ảo tại nhà để tiết kiệm thời gian đợi ở bệnh viện,...

* **Trong lĩnh vực vạn vật kết nối**

Các trợ lý ảo trên điện thoại thông minh cho phép người sử dụng tương tác, điều khiển các thiết bị trong mạng lưới vạn vật kết nối (Internet of Things - IoT), đặc biệt là công nghệ nhà thông minh (Smart Home). Trong hiện tương lai, công nghệ nhận dạng có thể giúp người dùng trực tiếp ‘ra lệnh’ cho các thiết bị này mà không cần thiết phải chạm trực tiếp vào điện thoại. Điều này tạo cảm giác thoải mái, tiện lợi cho người dùng.

## **1.1.3 Tiềm năng kinh tế cao**

Nhận dạng giọng nói cho phép chuyển đổi giọng nói thành văn bản, giúp việc tạo và sử dụng thông tin trở nên dễ dàng hơn bởi vì văn bản dễ dàng hơn để lưu trữ, xử lý và sử dụng, cho cả máy tính và con người. Các ứng dụng được xây dựng để phục vụ nhu cầu nhận dạng và tích hợp các dịch vụ nhận dạng đang ngày càng phát triển và là lĩnh vực hấp dẫn đối với thị trường công nghệ. Sẽ rất thuận tiện nếu chúng ta có thể sử dụng phần mềm dựa trên nhận dạng giọng nói để mua hàng, kiểm tra thời tiết, gửi email, tìm kiếm thông tin trên internet theo một cách mới để tương tác với máy móc.

Các ứng dụng chụp ảnh giải trí trên các thiết bị di động không những mang lại giá trị cho bản thân người dùng mà còn đem đến những giá trị về kinh tế cho phía nhà phát triển. Bắt kịp xu thế và nhu cầu của người dùng, các doanh nghiệp, cùng với những cải tiến ấn tượng về khả năng hiểu ngôn ngữ tự nhiên và tỷ lệ chính xác của nhận dạng giọng nói, đã khiến các công ty ngày càng muốn tiếp tục xây dựng tính năng hỗ trợ giọng nói trải nghiệm thậm chí vượt ra ngoài phạm vi riêng tư, các công ty, doanh nghiệp chủ động hơn trong cách tiếp cận với nhận dạng giọng nói để giới thiệu hoặc kết hợp công nghệ nhận dạng giọng nói vào các sản phẩm của họ.

Hơn thế nữa, lĩnh vực nhận dạng giọng nói và tổng quan hơn là trí tuệ nhân tạo sẽ ngày càng trở nên phức tạp hơn trong tương lai. Ngành công nghiệp này đã và đang thu hút với hàng trăm công ty trên thế giới thử nghiệm và tích hợp sản phẩm và dịch vụ của họ với công nghệ nhận dạng giọng nói. Theo lời của Brian Roemmele, người sáng lập và Tổng biên tập của tạp chí Multiplex, "60 năm qua, con người đã thích nghi với máy tính. 60 năm tới, máy tính sẽ thích ứng với chúng ta. Tiếng nói của chúng ta sẽ là nền tảng cho điều đó, nó sẽ là một cuộc cách mạng và nó sẽ thay đổi mọi thứ. ".

Trên cơ sở đó, việc xây dựng các dịch vụ nhận dạng giọng nói cũng như các ứng dụng các dịch vụ này vào các phần mềm hiện có trong nhiều lĩnh vực của cuộc sống đem lại một tiềm năng kinh tế rất cao.

## **1.1.4 Phục vụ nhu cầu sử dụng công nghệ cho đa dạng người trong xã hội**

Công nghệ ngày càng hiện đại, một số tầng lớp của xã hội gần như bị ‘lạc hậu’ đi vì khó khăn trong việc sử dụng các thiết bị điện tử thông minh. Người lớn tuổi, người cận thị, người tiếp cận với thiết bị điện tử hiện đại muộn khó khăn trong việc nhập thông tin từ bàn phím, màn hình điện thoại di động với phím nhỏ; người khiếm thị khó khăn trong việc sử dụng các thiết bị thông minh,... Công nghệ nhận dạng giọng nói giúp giải quyết các khó khăn trên.

# **1.2 LÝ DO LỰA CHỌN ĐỀ TÀI**

Trong những năm gần đây, các nhà nghiên cứu đang tập trung vào công nghệ nhận dạng giọng nói và đã có một số thành công đối với việc nhận dạng tiếng Anh và một số ngôn ngữ khác. Triển khai, xây dựng mô hình và ứng dụng các mô hình này vào thực tế ứng dụng vấn đề này là một việc làm hết sức có ý nghĩa đặc biệt trong giai đoạn công nghiệp hoá hiện đại hoá hiện nay.

Đối với Tiếng Việt, Tiếng Việt là một trong ngôn ngữ có nhiều thanh điệu nhất thế giới, cùng với những đặc trưng riêng về vùng miền, mặc dù đã có một số công trình nghiên cứu về lĩnh vực nhận dạng giọng nói Tiếng Việt nhưng lĩnh vực này nói chung vẫn còn khá mới ở nước ta.

Áp dụng các kiến thức đã học của bản thân trong quá trình học tập cũng như quá trình tìm hiểu, nghiên cứu vào việc xây dựng mô hình có ý nghĩa lớn, có tiềm năng cao trong tương lai, nhóm sinh viên lựa chọn đề tài “Xây dựng mô hình nhận dạng âm thanh tiếng Việt”. Mục tiêu cơ bản của đề tài này, nhóm sinh viên muốn nghiên cứu, thu thập dữ liệu và xây dựng mô hình nhận dạng âm thanh Tiếng Việt hiệu quả, đạt độ chính xác cao so với các ứng dụng hiện có trên thị trường.

Ngoài ra, việc chọn đề tài này giúp nhóm sinh viên tiếp cận với lĩnh vực học máy, nghiên cứu tìm hiểu thông tin từ các nguồn tài liệu quý giá. Hơn thế nữa, sau khi thực hiện đề tài, nhóm sinh viên sẽ có thêm kinh nghiệm và hiểu biết về quy trình làm ra một dự án thực tế, không những thế chúng em còn được học hỏi các kiến thức chuyên môn liên quan đến học sâu, huấn luyện mô hình, và áp dụng vào xây dựng ứng dụng minh họa. Chính những điều đó sẽ là nền tảng quý báu hỗ trợ đắc lực cho chúng em trên con đường học vấn và việc làm trong tương lai.

# **1.3 HƯỚNG PHÁT TRIỂN CỦA LUẬN VĂN**

Tiếng Việt được coi là một ngôn ngữ khó học với người nước ngoài bởi ngữ pháp, thanh điệu và đặc trưng vùng miền. Máy tính cũng giống như người nước ngoài - để nó nghe hiểu và diễn giải được giọng nói tiếng Việt thành dạng văn bản không phải là việc dễ dàng. Nhận dạng tiếng nói đóng vai trò quan trọng trong giao tiếp giữa người và máy. Nó giúp máy móc hiểu và thực hiện các hiệu lệnh của con người. Hiện nay trên thế giới, lĩnh vực nhận dạng tiếng nói đã đạt được nhiều tiến bộ vượt bậc. Đối với ngôn ngữ tiếng Anh, việc nhận dạng có thể đạt độ chính xác tới 99%. Ở Việt Nam, lĩnh vực nhận dạng giọng nói còn khá mới và hiện độ chính xác nhìn chung chưa cao. Luận văn này, nhóm sinh viên sẽ xây dựng, huấn luyện mô hình nhận dạng âm thanh Tiếng Việt và xây dựng ứng dụng demo với mục tiêu cơ bản là:

* Mô hình được xây dựng đạt độ chính xác tối thiểu 75%.
* Ứng dụng mẫu sẽ được phát triển trên nền tảng web với chức năng chính nhận một tập tin âm thanh Tiếng Việt và chuyển sang văn bản Tiếng Việt, bên cạnh đó cho phép ghi âm trực tiếp âm thanh và tải lên tập tin âm thanh, văn bản sau khi chuyển đổi hỗ trợ tải về ở nhiều định dạng.

# **1.4 MỤC TIÊU CỦA LUẬN VĂN**

Để hoàn thành tốt đề tài luận văn, bản luận văn và sản phẩm cuối cùng của nhóm sinh viên sẽ đảm bảo các mục tiêu tối thiểu sau đây:

* Bản luận văn trình bày lý thuyết nền tảng và giải pháp để xử lý việc nhận 1 tập tin âm thanh tiếng Việt và xuất ra nội dung văn bản ở dạng tiếng Việt.
* Xây dựng, thu thập dữ liệu, và đào tạo mô hình để nhận 1 file âm thanh tiếng Việt và xuất ra nội dung ở dạng văn bản
* Cải tiến độ chính xác của mô hình với mục tiêu là 75%.
* Xây dựng ứng dụng web chuyển đổi từ giọng nói sang văn bản Tiếng Việt. Ứng dụng mẫu áp dụng mô hình được xây dựng là ứng dụng web dựa trên nền tảng Flask. Ứng dụng cho phép định dạng, chỉnh sửa văn bản trực tuyến (căn lề, kích thước chữ, phông chữ, định dạng chữ, …), bên cạnh đó ứng dụng cho phép người dùng tải lên tập tin âm thanh ít nhất 3 định dạng (.mp3, .acc, .wav). và cho phép người dùng tải xuống tập tin văn bản ít nhất 3 định dạng (.txt, .pdf, .doc). Ứng dụng cũng hỗ trợ ghi âm, nghe lại đoạn âm thanh đã ghi âm trước khi thực hiện chuyển đổi.

# **1.5 PHẠM VI ĐỀ TÀI**

Sản phẩm của đề tài “Xây dựng mô hình nhận dạng âm thanh tiếng Việt” là một mô hình cho phép nhận vào một tập tin âm thanh Tiếng Việt và trả về một đoạn văn bản Tiếng Việt. Nhóm sinh viên sẽ tập trung vào việc phát triển một cách đầy mô hình này dựa trên engine đã có sẵn trên thị trường. Khả năng nhận dạng ngôn ngữ Tiếng Việt của mô hình được đào tạo từ bộ dữ liệu âm thanh Tiếng Việt do nhóm tìm hiểu, thu thập từ các nguồn dữ liệu mở. Sản phẩm luận văn được thực hiện một cách toàn diện nên sẽ được áp dụng, mở rộng vào các dịch vụ web (API). Nhóm sẽ xây dựng một API sử dụng mô hình trên và một trang web mẫu sử dụng API này để biểu diễn mô hình.

# **TÀI LIỆU THAM KHẢO**