

ĐOÀN TNCS HỒ CHÍ MINH THÀNH PHỐ HÀ NỘI
ĐOÀN THANH NIÊN ĐẠI HỌC BÁCH KHOA HÀ NỘI
ĐOÀN THANH NIÊN TRƯỜNG ĐẠI HỌC NGOẠI THƯƠNG



ĐỀ ÁN SẢN PHẨM CÔNG NGHỆ
CHƯƠNG TRÌNH ƯƠM TẠO CÔNG NGHỆ
BUSITECH BOOTCAMP

VÒNG 3 - INTENSIVE BOOTCAMP

Tên đội tham dự:

S.A.T.I.S

Lĩnh vực:

AI ứng dụng

Tên đề tài:

**TalkTrack: Cải thiện chất lượng
dịch vụ chăm sóc khách hàng**

Hà Nội, tháng 4, 2023

Lời cam kết

Chúng tôi xin cam đoan sản phẩm/giải pháp là công trình nghiên cứu của riêng tập thể chúng tôi, không sao chép của ai, do chúng tôi tự nghiên cứu, đọc, dịch tài liệu, tổng hợp và thực hiện. Nội dung trong đề án tôi có sử dụng một số tài liệu tham khảo trên Internet nhưng không sao chép hoàn toàn. Các số liệu, chương trình phần mềm và những kết quả trong đề án này là trung thực và được thực hiện bởi chính bản thân chúng tôi.

Chúng tôi xin cam kết chịu trách nhiệm trước pháp luật, trước Ban Tổ chức Chương trình và chấp nhận mọi hình thức cảnh cáo, kỷ luật nếu xảy ra bất kì hiện tượng gian lận, đạo văn nào trong toàn bộ đề án này.

TM. Đội tham dự
Đại diện

Ngô Thị Thu Huyền

Mục lục

Lời cam kết	2
Mục lục	3
1 Vấn đề thực tiễn và giải pháp tổng quan	5
1.1 Vấn đề thực tiễn.	5
1.2 Giải pháp đề ra.	5
2 Tổng quan đề án	7
2.1 Mô tả chung.	7
2.1.1 Tổng quan.	7
2.1.2 Phân tích SWOT.	8
2.1.3 BMC.	8
2.2 Phân tích thị trường.	10
2.2.1 Thị trường mục tiêu.	10
2.2.2 Thị phần và giá trị thị trường.	10
2.2.3 Đối thủ cạnh tranh.	11
2.2.4 Ưu thế so sánh.	12
2.3 Chiến lược tiếp thị.	13
2.3.1 Chiến lược tiếp thị.	13
2.3.2 Chiến lược bán hàng.	14
3 Mô tả kỹ thuật	17
3.1 Công nghệ sử dụng.	17
3.2 Bài toán kỹ thuật.	17
3.2.1 Bài toán kỹ thuật số 1: Nhận diện cảm xúc người nói.	18
3.2.2 Bài toán kỹ thuật số 2: Chuyển đổi giọng nói thành văn bản.	18
3.2.3 Bài toán kỹ thuật số 3: Trích xuất từ khóa.	18
3.2.4 Bài toán kỹ thuật số 4: Trích xuất từ ngữ thô tục.	18
3.2.5 Bài toán kỹ thuật số 5: Thiết lập hệ thống tự động.	19
3.3 Hướng giải quyết bài toán.	19
3.3.1 Hướng giải quyết bài toán số 1: Nhận diện cảm xúc người nói.	19
Đề án Sản phẩm Công nghệ - Lĩnh vực AI ứng dụng	3

3.3.2	Hướng giải quyết bài toán số 2: Chuyển đổi giọng nói thành văn bản.	19
3.3.3	Hướng giải quyết bài toán số 3 và số 4: Trích xuất từ khóa và trích xuất từ ngữ thô tục.	19
3.3.4	Hướng giải quyết bài toán số 5: Thiết lập hệ thống tự động.	20
3.4	Mô tả chi tiết sản phẩm.	20
3.4.1	Tính năng số 1: Đánh giá chất lượng cuộc gọi.	20
3.4.2	Tính năng số 2: Phân tích tổng quan.	20
3.4.3	Tính năng số 3: Đào tạo nghiệp vụ chăm sóc khách hàng.	21
3.5	Phân tích chuyên sâu sản phẩm.	21
3.5.1	Tính đổi mới, tối ưu trong công nghệ.	21
3.5.2	Trải nghiệm người dùng.	21
3.5.3	Tiềm năng phát triển rộng rãi của sản phẩm.	26
4	Nguồn lực	27
4.1	Đội ngũ nhân sự.	27
4.2	Nguồn lực nhân sự khác.	28
4.3	Nguồn lực tài chính và cơ sở hạ tầng khác.	28

Chương 1

Vấn đề thực tiễn và giải pháp tổng quan

1.1 Vấn đề thực tiễn.

Hiện nay, các doanh nghiệp đang ngày càng chú trọng vào yếu tố trải nghiệm của người dùng khi sử dụng dịch vụ. Tuy nhiên, đối với loại hình chăm sóc khách hàng qua điện thoại, các doanh nghiệp chưa có giải pháp giúp tự động hóa quy trình quản lý thái độ, chất lượng của nhân viên hỗ trợ người dùng. Hiện nay, các doanh nghiệp vẫn đang sử dụng phương pháp đánh giá chất lượng thủ công bằng cách chọn ngẫu nhiên từ 1% - 3% cuộc gọi để kiểm tra lại, điều này dẫn đến việc đánh giá không được toàn diện, chính xác cũng như trở nên tốn chi phí, khó khăn đối với các doanh nghiệp có số lượng cuộc gọi rất lớn. Ngoài ra, việc đào tạo nhân viên CSKH hiện nay cũng đang được diễn ra thủ công, thiếu giải pháp giúp tối ưu việc cải thiện chất lượng học viên.

1.2 Giải pháp đề ra.

Chúng tôi đề xuất giải pháp sử dụng trí tuệ nhân tạo để tự động hóa quy trình quản lý chất lượng dịch vụ chăm sóc khách hàng của các doanh nghiệp. Cuộc gọi giữa nhân viên CSKH và khách hàng sẽ được ghi lại, đưa vào mô hình trí tuệ nhân tạo để phát hiện những ngôn từ tiêu cực, cũng như đưa ra đánh giá về cảm xúc, thái độ của nhân viên trong quá trình hỗ trợ khách hàng qua điện thoại. Bên cạnh đó, quy trình đào tạo nhân viên CSKH cũng được tự động hóa. Toàn bộ hệ thống được thiết kế để bảo mật dữ liệu về doanh nghiệp, người dùng, đảm bảo dữ liệu của khách hàng được sử dụng đúng mục đích.

Chương 2

Tổng quan đề án

2.1 Mô tả chung.

2.1.1 Tổng quan.

TalkTrack là một công cụ sử dụng trí tuệ nhân tạo nhằm để giám sát thái độ chăm sóc khách hàng của nhân viên cũng như trạng thái cảm xúc của khách hàng khi trải nghiệm dịch vụ, đồng thời cung cấp hệ thống đào tạo thái độ nhân viên một cách tự động. TalkTrack dựa trên các câu đối thoại giữa nhân viên dịch vụ và khách hàng để đưa ra thái độ của người nói, nhằm để xác định các câu thoại mang trạng thái như tích cực, tiêu cực... Từ đó, các doanh nghiệp, tổ chức có thể phát hiện ra những câu nói không chuẩn mực của nhân viên, hay thời điểm khách hàng cảm thấy không vừa lòng, để từ đó có thể dễ dàng giải quyết vấn đề và đưa ra định hướng cải thiện chất lượng dịch vụ.

TalkTrack được xây dựng bởi hai mô hình chính: ASR (Automatic Speech Recognition) dùng để nhận dạng giọng nói và chuyển qua dạng chữ viết và SER (Speech Emotion Recognition) để phân tích cảm xúc người nói. TalkTrack có thể được lồng ghép với môi trường đàm thoại giữa khách hàng và nhân viên, sau đó phân tích những câu thoại trong quá trình đối thoại giữa hai bên, từ đó có thể giám sát thái độ làm việc của nhân viên, cũng như đánh giá cảm xúc của khách hàng trong suốt quá trình đối thoại. Quá trình giám sát được diễn ra một cách tự động và liên tục giúp cho các doanh nghiệp giảm chi phí nhân lực và tạo một môi trường dịch vụ thân thiện với khách hàng.

TalkTrack sẽ là một công cụ hữu hiệu để cho các doanh nghiệp đào tạo nhân viên, cải thiện dịch vụ chăm sóc khách hàng. Đặc biệt, TalkTrack sẽ là một trợ thủ đắc lực cho các doanh nghiệp đòi hỏi nhiều sự trao đổi giữa khách hàng với nhân viên. Điều mà TalkTrack hướng đến chính là một môi trường dịch vụ thân thiện và tích cực.

2.1.2 Phân tích SWOT.

Điểm mạnh	Điểm yếu
<ul style="list-style-type: none"> - Sản phẩm có sự đổi mới so với đối thủ cạnh tranh, đó là đưa quá trình đào tạo dịch vụ chăm sóc khách hàng một cách tự động. - Tự chủ trong phát triển công nghệ AI, đặc biệt là xử lý tiếng nói tiếng Việt, trong khi nguồn dữ liệu bằng Tiếng Việt là rất hạn chế. 	<ul style="list-style-type: none"> - Nhóm thiếu kinh nghiệm triển khai các mô hình, cũng như thiếu nguồn lực về vốn, cơ sở hạ tầng thiết bị để triển khai sản phẩm trên thực tế. - Các thành viên nhóm chưa được trang bị bài bản, đầy đủ được về các kiến thức kinh tế, thị trường.
Cơ hội	Thách thức
<ul style="list-style-type: none"> - Các ngành dịch vụ đòi hỏi trao đổi giữa nhân viên tư vấn và khách hàng ngày càng nhiều (tư vấn y tế, giao dịch ngân hàng, hỏi đáp thắc mắc về ví điện tử,...). Điều đó đã thúc đẩy mạnh nhu cầu về công cụ giám sát thái độ nhân viên, cũng như hệ thống đào tạo nhân viên CSKH một cách hoàn toàn tự động. - Công cụ để phân tích và giám sát thái độ chăm sóc khách hàng của nhân viên có rất ít, hiện tại chỉ có hai đối thủ thực hiện điều này (InterITS và EM&AI) tuy nhiên chưa có sự cải tiến để phù hợp với xu hướng hiện tại. 	<ul style="list-style-type: none"> - Dữ liệu dùng để huấn luyện bằng Tiếng Việt còn rất hạn chế, phải tự tìm và trích xuất. - Câu từ, ngữ pháp Tiếng Việt rộng lớn và khá phức tạp. - Đối thủ có tiềm lực cũng như kinh nghiệm triển khai sản phẩm trên thực tế, ngoài ra họ cũng sở hữu nguồn dữ liệu lớn để huấn luyện cho mô hình.

2.1.3 BMC.

Phân khúc khách hàng (Customer Segments):

- Theo quy mô doanh nghiệp, TalkTrack hướng đến phân khúc khách hàng quy mô vừa. Đây là những doanh nghiệp có chú trọng kiểm soát và cải thiện dịch vụ CSKH.

Phương án giá trị (Value Propositions):

- Cung cấp dịch vụ đánh giá chất lượng của nhân viên chăm sóc khách hàng, giải quyết vấn đề đánh giá thủ công tiêu tốn nhiều thời gian.
- Cung cấp dịch vụ đào tạo nhân viên chăm sóc khách hàng một cách tự động.

- Nâng cao chất lượng dịch vụ trong thời kỳ cạnh tranh cao và khách hàng ngày càng kỹ tính.

Kênh cung cấp (Channels):

- Website của dự án, cung cấp đầy đủ thông tin và có bản dùng thử.
- Các hội thảo ứng dụng trí tuệ nhân tạo, đặc biệt là xử lý ngôn ngữ tự nhiên.

Quan hệ khách hàng (Customer Relationships):

- Hỗ trợ kỹ thuật và tư vấn sản phẩm cho khách hàng giúp khách hàng dễ dàng tiếp cận.
- Khảo sát ý kiến khách hàng, sẵn sàng thay đổi, nâng cấp sản phẩm phục vụ nhu cầu khách hàng.

Dòng doanh thu (Revenue Stream):

- Doanh thu trực tiếp từ quá trình bán các gói dịch vụ sản phẩm như gói đánh giá chất lượng nhân viên, gói đào tạo tự động nhân viên.

Nguồn tiềm lực chính (Key Resources):

- Các công nghệ, thuật toán AI tân tiến nhất về chuyển đổi giọng nói thành văn bản, nhận dạng cảm xúc.
- Đội ngũ nhân sự có chuyên môn cao.

Công việc chính (Key Activities):

- Kiểm soát, giám sát và đào tạo khả năng chăm sóc khách hàng trực tiếp hoặc qua điện thoại của nhân viên các doanh nghiệp.
- Xây dựng phát triển các thuật toán AI để cải thiện độ chính xác của sản phẩm.
- Hỗ trợ tư vấn khách hàng.
- Nghiên cứu, phát triển thêm tính năng cho sản phẩm.

Đối tác chính (Key Partnerships):

- Các nhà cung cấp dịch vụ điện toán đám mây (Cloud Computing).
- Các công ty cung cấp và tư vấn giải pháp CRM (Customer Relationship Management).
- Các doanh nghiệp cần cải thiện dịch vụ chăm sóc khách hàng.

Cơ cấu chi phí (Cost Structure):

- Nghiên cứu và phát triển công nghệ, hệ thống (RnD).

- Tiếp thị, tìm kiếm khách hàng tiềm năng.
- Thu thập và xử lý dữ liệu.
- Duy trì, mở rộng cơ sở hạ tầng công nghệ (Cloud, Server).

2.2 Phân tích thị trường.

2.2.1 Thị trường mục tiêu.

Thị trường mục tiêu: Việt Nam.

Khách hàng mục tiêu: Các doanh nghiệp tài chính ngân hàng. Đây là các doanh nghiệp luôn mà chất lượng dịch vụ của họ ảnh hưởng trực tiếp đến doanh thu.

Tại thị trường Việt Nam, các sản phẩm sử dụng công nghệ xử lý ngôn ngữ tự nhiên còn chưa phổ biến và có nhiều hạn chế. Bên cạnh đó, điều kiện sống của người Việt Nam ngày càng cải thiện, khách hàng ngày càng khó tính, dẫn đến nhu cầu về chăm sóc khách hàng đối với các doanh nghiệp. Theo một khảo sát của Vietnamnet Media Group :

- Hơn 60% người tiêu dùng ở Việt Nam tin rằng các doanh nghiệp nên tập trung vào việc cải thiện dịch vụ khách hàng để nâng cao uy tín thương hiệu.
- 74% công ty nói rằng việc chuyển đổi khách hàng tiềm năng thành khách hàng là ưu tiên hàng đầu của họ.
- 65% doanh nghiệp áp dụng công nghệ vào quy trình CSKH trong vòng 5 năm đầu tiên kể từ khi ra mắt.

Bởi vậy thị trường Việt Nam là một thị trường rất tiềm năng cho sản phẩm.

TalkTrack hướng đến phân khúc dành cho các ngân hàng và các ví điện tử, có quy mô trung bình nhưng số lượng nhân viên chăm sóc khách hàng cần kiểm soát, đòi hỏi giải pháp tự động.

2.2.2 Thị phần và giá trị thị trường.

Giá trị tiềm năng: Tới thời điểm hiện tại, ngày càng nhiều mô hình dịch vụ diễn ra hàng ngày, việc trao đổi giữa chuyên viên tư vấn chăm sóc khách hàng và người sử dụng dịch vụ ngày càng nhiều. Theo thống kê của VINNO (<http://vinno.vn>), vào năm 2021, người ta ước tính rằng các công ty đã chi 12% ngân sách tiếp thị của họ cho phần mềm CSKH. TalkTrack luôn là một sự lựa chọn phù hợp cho thị trường dịch vụ hiện nay. Một số thị trường tiềm năng của TalkTrack:

- Doanh nghiệp ngân hàng, tài chính, bảo hiểm.
- Dịch vụ y tế, chăm sóc sức khỏe.

- Dịch vụ tư vấn làm đẹp, thẩm mỹ viện.

Trước mắt, TalkTrack tập trung vào nhóm khách hàng là các doanh nghiệp ngân hàng và các ví điện tử. Theo báo cáo của Công ty nghiên cứu thị trường Global Data, thị trường CSKH trong lĩnh vực ngân hàng tại Việt Nam đang có mức tăng trưởng ấn tượng, dự kiến sẽ đạt tỷ lệ tăng trưởng hàng năm trung bình khoảng 14% trong giai đoạn 2019-2024.

Thị phần: Với hai đối thủ cạnh tranh là InterITS và EM&AI, nhưng với những đặc điểm cải tiến, TalkTrack ước tính có thể chiếm tới 20% thị phần công nghệ AI.

2.2.3 Đối thủ cạnh tranh.

Dưới đây là bảng so sánh các đối thủ cạnh tranh trực tiếp với sản phẩm TalkTrack theo nhiều tiêu chí khác nhau.

Công ty đối thủ	Inter ITS	EM&AI
Sản phẩm cạnh tranh	Emotional Assistant	Virtual QC
Tính năng	<ul style="list-style-type: none"> - Chuyển đổi giọng nói thành văn bản (Automatic Speech Recognition). - Nhận dạng cảm xúc người nói (Speech Emotion Recognition). - Giám sát chất lượng cuộc gọi. - Đánh giá và xử lý cuộc gọi. - Tích hợp với các sản phẩm tổng đài. 	<ul style="list-style-type: none"> - Chuyển đổi giọng nói thành văn bản (Automatic Speech Recognition). - Nhận dạng cảm xúc người nói (Speech Emotion Recognition). - Giám sát chất lượng cuộc gọi. - Đánh giá và xử lý cuộc gọi. - Tích hợp với các sản phẩm tổng đài. - Báo cáo và phân tích thái độ nhân viên. - Bình luận và đào tạo cho tổng đài viên.
Giá thành	Không công bố.	Không công bố.
Tệp khách hàng	Các doanh nghiệp.	Các doanh nghiệp.
Điểm mạnh	<ul style="list-style-type: none"> - Hệ thống có áp dụng phân tích ngữ nghĩa nội dung để bổ sung đánh giá thái độ nhân viên. 	<ul style="list-style-type: none"> - Hệ thống có thể được tùy chỉnh theo nhu cầu kinh doanh. - Giao diện trực quan, dễ sử dụng.
Điểm yếu	<ul style="list-style-type: none"> - Không có báo cáo và phân tích thái độ nhân viên. - Không có bình luận và đào tạo nhân viên CSKH. 	<ul style="list-style-type: none"> - Quy trình đào tạo nhân viên là hoàn toàn thủ công.

2.2.4 Ưu thế so sánh.

Với công nghệ mới nổi bật, đó là đưa quá trình đào tạo dịch vụ chăm sóc khách hàng một cách tự động, TalkTrack mang nhiều lợi ích so với các đối thủ cạnh tranh. Việc đào tạo tự động sẽ giúp dịch vụ chăm sóc khách hàng được cải thiện một cách hiệu quả ngay cả với các nhân viên mới. Thời gian và công sức cho việc đào tạo giảm đi đáng kể. Đây là công nghệ chưa được các đối thủ cạnh tranh tận dụng.

TalkTrack có thể được tùy biến linh hoạt dựa trên nhu cầu của khách hàng. Ngoài việc có

thể đưa ra đánh giá các audio giọng nói như tiêu cực, tích cực... TalkTrack có thể chấm điểm theo thang điểm mà khách hàng mong muốn, hoặc cho ra kết quả ở định dạng nhãn khác như vui vẻ, tức giận... Điều đó tùy thuộc vào nhu cầu của khách hàng.

2.3 Chiến lược tiếp thị.

2.3.1 Chiến lược tiếp thị.

Đối tượng mục tiêu của chúng tôi những doanh nghiệp tài chính ngân hàng. Chúng tôi sử dụng mô hình 4P để mô hình hóa chiến lược tiếp thị của sản phẩm.

- Về sản phẩm (Product): TalkTrack được hình thành dựa trên nhu cầu nâng cao chất lượng dịch vụ ở các doanh nghiệp sử dụng dịch vụ CSKH. Khả năng giám sát thái độ và đào tạo nhân viên của TalkTrack hoàn toàn có thể thỏa mãn nhu cầu giải quyết các vấn đề liên quan đến thái độ đối xử với khách hàng của các nhân viên trong doanh nghiệp. Việc giám sát lời nói và đào tạo nhân viên đều diễn ra một cách tự động, và đều được định hướng sao cho phù hợp với tính chất của các doanh nghiệp. TalkTrack là sản phẩm dễ sử dụng, thân thiện với người dùng và đạt được độ chính xác cao.

- Về địa điểm (Places): TalkTrack sẽ sử dụng kênh phân phối trực tiếp. Địa điểm trưng bày và mua bán sản phẩm bao gồm: trên website chính của TalkTrack, các hội thảo về công nghệ trí tuệ nhân tạo.

- Về giá (Price): Chúng tôi đóng gói sản phẩm để có thể cung cấp 5 loại gói sản phẩm chính, gói cơ bản sẽ sử dụng một số tính năng với một thời gian giới hạn, gói nâng cấp sẽ giúp người dùng truy cập đầy đủ các tính năng với trải nghiệm tốt nhất, mỗi gói tùy thuộc nhu cầu người dùng như khóa premium về đào tạo nhân viên CSKH, gói Premium Training, gói Premium QC. Ngoài ra, chúng tôi cũng đưa ra những gói dùng thử 7 ngày trong thời kỳ đầu để lấy được thị phần từ đối thủ. Khi đạt được một số lượng khách hàng nhất định, dữ liệu thu được từ quá trình chạy thực tế được sử dụng để cải tiến mô hình, giúp tăng độ chính xác, tin cậy, từ đó tạo ra lợi thế cạnh tranh đối với các đối thủ.

- Về Quảng bá (Promotion): Chúng tôi chủ yếu sử dụng email marketing để quảng bá sản phẩm tới các doanh nghiệp như ngân hàng, thẩm mỹ viện, ... Ngoài ra, việc đẩy mạnh truyền thông qua các mạng xã hội (Facebook, Tiktok) với mục đích truyền bá sản phẩm giúp tăng lượng truy cập vào website, tăng lượng người dùng đăng ký và dùng thử sản phẩm. Bởi TalkTrack là một sản phẩm công nghệ ứng dụng AI, cụ thể là ứng dụng NLP - Xử lý ngôn ngữ tự nhiên nên sẽ gây tò mò và hứng thú cho nhiều người. Tạo các cộng đồng về các sản phẩm công nghệ AI, về những người quan tâm đến công nghệ AI, chia sẻ các kiến thức về công nghệ AI thông qua các hội thảo AI.

2.3.2 Chiến lược bán hàng.

Để bán sản phẩm cho thị trường mục tiêu, chúng tôi áp dụng các chiến lược như sau:

- **Kênh bán:** sản phẩm được trưng bày online trên website và bán trực tiếp qua hình thức kí kết hợp đồng.
- **Chiến lược hợp tác:** có thể hợp tác với các công ty chuyên cung cấp và tư vấn giải pháp CRM (Customer relationship management) . Talktrack có thể tích hợp vào và cải thiện phần mềm CRM của họ.
- **Chính sách mua bán:**
 - Khách hàng trước khi thực hiện mua bán sẽ được thêm thông tin về sản phẩm. bao gồm: các bài viết, các công nghệ sử dụng, hướng dẫn sử dụng, trải nghiệm miễn phí... để tạo sự tin tưởng.
 - Khách hàng sau khi mua sản phẩm sẽ được cung cấp các dịch vụ hỗ trợ khách hàng chuyên nghiệp, như trợ giúp trực tuyến, hỗ trợ qua email hoặc điện thoại để giải đáp mọi thắc mắc của khách hàng về sản phẩm.
- **Chiến lược mua bán:** Sản phẩm sẽ được chia thành các gói để cho khách hàng có thể xem xét và lựa chọn gói phù hợp.

Loại gói	Nội dung
Gói Trial	<ul style="list-style-type: none">- Hoàn toàn miễn phí, chỉ được sử dụng một lần trên một tài khoản.- Trải nghiệm trong 7 ngày.- Được dùng tất cả các tính năng.- Tổng giới hạn thời lượng gọi là 30 phút.
Gói Standard	<ul style="list-style-type: none">- Cung cấp tính năng đánh giá chất lượng.- Tổng giới hạn thời lượng 60 tiếng một tháng
Gói Standard Training	<ul style="list-style-type: none">- Cung cấp tính năng đào tạo nhân viên- Tổng giới hạn số bài luyện tập là 5000 bài/1 ngày.
Gói Premium QC	<ul style="list-style-type: none">- Cung cấp tính năng đánh giá chất lượng.- Tổng giới hạn thời lượng 120 tiếng một tháng.
Gói Premium Training	<ul style="list-style-type: none">- Cung cấp tính năng đào tạo nhân viên.- Tổng giới hạn số bài luyện tập là 10000/1 ngày.

Ngoài ra, khách hàng khi sử dụng nhiều hơn hai gói sản phẩm sẽ được hưởng ưu đãi về giá cả cũng như thời hạn sử dụng sản phẩm:



- Khi mua từ hai gói sản phẩm trở lên thì sẽ được khuyến mãi 10
- Khi mua gói bất kỳ với thời hạn 6 tháng trở lên thì sẽ được gia hạn thêm một tháng sử dụng đối với sản phẩm đó.

Chương 3

Mô tả kỹ thuật

3.1 Công nghệ sử dụng.

Dưới đây là danh sách các công nghệ mà chúng tôi sử dụng để hoàn thiện sản phẩm TalkTrack. Ngoài những công nghệ lõi để phát triển mô hình AI, chúng tôi còn áp dụng các công nghệ mới giúp hỗ trợ việc triển khai sản phẩm demo:

- **HuggingFace Transformers** - sử dụng trong việc huấn luyện mô hình chuyển đổi giọng nói thành văn bản (Automatic Speech Recognition), trích xuất từ khóa (Keyword Extraction) và trích xuất từ ngữ thô tục (Offensive Keyword Extraction).
- **SpeechBrain** - sử dụng trong việc tích hợp mô hình nhận dạng cảm xúc trong giọng nói (Speech Emotion Recognition).
- **ONNX - Open Neural Network Exchange** - sử dụng trong việc chuyển đổi các mô hình AI về các dạng cố định, giúp tăng tốc độ xử lý của các mô hình.
- **FastAPI** - sử dụng trong việc xây dựng API tích hợp 2 mô hình AI nói trên.
- **ngrok** - sử dụng trong việc triển khai API, giúp API có thể được truy cập từ các trang web.
- **Github Pages** - sử dụng trong việc triển khai sản phẩm Web demo.

3.2 Bài toán kỹ thuật.

TalkTrack là một sản phẩm công nghệ tích hợp AI, cụ thể là xử lý ngôn ngữ tự nhiên. Từ những vấn đề thực tiễn, ba bài toán kỹ thuật sau đây được đề ra để hoàn thiện TalkTrack:

- Nhận diện cảm xúc người nói.
- Chuyển đổi giọng nói thành văn bản.

- Trích xuất từ khóa.
- Trích xuất từ ngữ thô tục.
- Thiết lập hệ thống tự động.

3.2.1 Bài toán kỹ thuật số 1: Nhận diện cảm xúc người nói.

Đây là công nghệ cốt lõi của sản phẩm - Speech Emotion Recognition (nhận diện cảm xúc người nói). Sản phẩm của chúng tôi có thể nhận diện thái độ trong giọng nói của khách hàng và nhân viên chăm sóc khách hàng, ngoài ra có thể chấm điểm chất lượng cuộc gọi hoặc người nói.

Ý tưởng chính của công nghệ là xây dựng một mô hình học sâu (mô hình chuyên sâu cho bài toán Speech Emotion Recognition) có thể được huấn luyện với dữ liệu giọng nói hội thoại tiếng Việt, đặc biệt là dữ liệu hội thoại liên quan đến chủ đề chăm sóc khách hàng. Sau khi training, quá trình đánh giá cảm xúc người nói sẽ hoàn toàn tự động.

3.2.2 Bài toán kỹ thuật số 2: Chuyển đổi giọng nói thành văn bản.

Đây là bài toán phụ của sản phẩm. Việc chuyển đổi giọng nói thành văn bản (bài toán Automatic Speech Recognition) giúp quá trình đào tạo nhân viên dễ dàng hơn, người quản lý dễ dàng xem xét các câu từ trong cuộc hội thoại mà không cần phải nghe thủ công.

Ý tưởng chính của công nghệ là xây dựng mô hình học sâu (mô hình chuyên sâu cho bài toán Automatic Speech Recognition) được huấn luyện trên các tập dữ liệu giọng nói tiếng Việt. Sau quá trình training, quá trình chuyển đổi giọng nói thành văn bản sẽ hoàn toàn tự động.

3.2.3 Bài toán kỹ thuật số 3: Trích xuất từ khóa.

Ngoài những công nghệ AI để xử lý giọng nói, sản phẩm TalkTrack cũng bao gồm các mô hình AI để xử lý văn bản. Trích xuất từ khóa (Keyword Extraction) là việc tìm kiếm những từ ngữ quan trọng trong một câu nói hay một đoạn hội thoại. Với sản phẩm TalkTrack, việc trích xuất từ khóa trong câu nhằm giúp cho các nhà quản lý có thể nhanh chóng nắm bắt được nội dung chính của một cuộc gọi mà không cần phải nghe hay xem hết nội dung văn bản của cuộc gọi đó.

3.2.4 Bài toán kỹ thuật số 4: Trích xuất từ ngữ thô tục.

Sản phẩm của chúng tôi được tích hợp công nghệ trích xuất từ ngữ thô tục (Offensive Keyword Extraction), nhằm phát hiện và phân tích các trường hợp có thể dẫn đến thái độ tiêu cực từ cả khách hàng và nhân viên. Ngoài ra, việc trích xuất từ ngữ thô tục cũng hỗ trợ cho việc đánh giá chất lượng nghiệp vụ của nhân viên CSKH, đảm bảo rằng nhân viên sử dụng ngôn từ phù hợp khi trao đổi với khách hàng.

3.2.5 Bài toán kỹ thuật số 5: Thiết lập hệ thống tự động.

Sau khi xây dựng mô hình giải quyết các vấn đề chính của sản phẩm, chúng tôi thiết lập hệ thống tự động trong việc đánh giá chất lượng dịch vụ và đào tạo nhân viên. Hệ thống này (hiện tại được xây dựng là một website) giúp người dùng dễ dàng sử dụng sản phẩm. Người dùng ở đây có thể là người quản lý cần kiểm soát chất lượng các nhân viên chăm sóc khách hàng hoặc có thể là những học viên hay nhân viên tham gia quá trình đào tạo tự động.

3.3 Hướng giải quyết bài toán.

3.3.1 Hướng giải quyết bài toán số 1: Nhận diện cảm xúc người nói.

Để giải quyết bài toán nhận diện cảm xúc người nói, chúng tôi sử dụng mô hình Wav2vec2 - một trong những mô hình Deep Learning tốt nhất cho việc xử lý giọng nói hiện nay. Hiện tại, do chưa thể xây dựng hay truy cập tới nguồn dữ liệu tiếng Việt dành riêng cho bài toán, chúng tôi quyết định sử dụng mô hình Wav2Vec2 đã được huấn luyện trên nhiều bộ dữ liệu từ nhiều ngôn ngữ khác nhau. Nhờ vậy, mô hình nhận diện cảm xúc vẫn đạt kết quả tốt khi xử lý tiếng nói tiếng Việt, bởi vì cảm xúc người nói không phụ thuộc quá nhiều vào ngôn ngữ. Mô hình nhận diện cảm xúc huấn luyện sẵn được thu thập từ framework SpeechBrain.

3.3.2 Hướng giải quyết bài toán số 2: Chuyển đổi giọng nói thành văn bản.

Với bài toán chuyển đổi giọng nói thành văn bản, việc thu thập và xử lý dữ liệu tiếng Việt dễ dàng hơn rất nhiều so với bài toán nhận diện cảm xúc người nói tiếng Việt. Vì vậy, chúng tôi đã huấn luyện một mô hình học sâu với một lượng lớn dữ liệu tiếng nói tiếng Việt, được chọn lựa kỹ lưỡng để đảm bảo tập dữ liệu đa dạng theo nhiều tiêu chí, ví dụ như giới tính, vùng miền và độ dài câu nói. Về kiến trúc của mô hình được sử dụng, chúng tôi chọn mô hình Wav2Vec2 để huấn luyện, tương tự kiến trúc đã dùng trong bài toán nhận diện cảm xúc. Mô hình chuyển đổi giọng nói thành văn bản sau khi được huấn luyện với nguồn dữ liệu lớn đã đạt kết quả rất tốt: tỷ lệ lỗi từ (Word Error Rate) đạt 5% - 10% trên các tập dữ liệu kiểm thử công khai dành cho tiếng Việt. Để huấn luyện mô hình, chúng tôi sử dụng framework HuggingFace Transformers.

3.3.3 Hướng giải quyết bài toán số 3 và số 4: Trích xuất từ khóa và trích xuất từ ngữ thô tục.

Để giải quyết bài toán trích xuất từ khóa và trích xuất từ ngữ thô tục, chúng tôi sử dụng mô hình RoBERTa được huấn luyện trước trên một lượng lớn dữ liệu không nhãn tiếng Việt. Sau đó, mô hình được bổ sung các lớp tuyến tính để có thể cùng lúc trích xuất từ khóa và từ ngữ thô tục xuất hiện trong câu nói, và được huấn luyện lại với dữ liệu tiếng Việt theo chủ đề

chăm sóc khách hàng để đảm bảo mô hình đạt độ chính xác cao khi xử lý dữ liệu cùng chủ đề.

3.3.4 Hướng giải quyết bài toán số 5: Thiết lập hệ thống tự động.

Việc đầu tiên cần làm trước khi thiết lập hệ thống là triển khai các mô hình AI thông qua một API. Để đảm bảo các mô hình AI đạt tốc độ xử lý cao mà không tốn quá nhiều tài nguyên, chúng tôi đã sử dụng framework ONNX để chuyển đổi các mô hình AI về các dạng cố định, từ đó tăng tốc độ xử lý của các mô hình.

Tiếp theo, để triển khai các mô hình AI, chúng tôi sử dụng FastAPI và ngrok. FastAPI được ứng dụng để lập trình API, cho phép tích hợp 2 mô hình AI nói trên. Với ngrok, chúng tôi sử dụng trong việc tạo một cổng (tunnel) từ máy chủ cá nhân (localhost) tới mạng Internet, giúp cho các trang web có được quyền truy cập tới API.

3.4 Mô tả chi tiết sản phẩm.

Sản phẩm TalkTrack cung cấp 3 tính năng chính cho khách hàng: Đánh giá chất lượng cuộc gọi, Phân tích tổng quan, Đào tạo nghiệp vụ chăm sóc khách hàng. Các phần dưới đây sẽ giải thích chi tiết các tính năng nêu trên.

3.4.1 Tính năng số 1: Đánh giá chất lượng cuộc gọi.

Để đánh giá chất lượng cuộc gọi, chúng tôi sử dụng 2 mô hình AI - nhận diện cảm xúc và chuyển đổi giọng nói thành văn bản. Với mỗi cuộc gọi, nhờ việc ứng dụng 2 mô hình AI nói trên, hệ thống sẽ đưa ra nội dung văn bản cũng như cảm xúc của từng câu nói (tích cực, tiêu cực, bình thường). Từ đó, các quản lý viên cũng như nhân viên chăm sóc khách hàng có thể đánh giá được chất lượng cuộc gọi.

Ngoài ra, để giúp quản lý viên nắm bắt nhanh chóng nội dung chính của cuộc đối thoại mà không cần nghe hay xem hết nội dung văn bản, hệ thống cũng đưa ra các từ khóa mang tính quan trọng của từng câu nói trong cuộc gọi.

3.4.2 Tính năng số 2: Phân tích tổng quan.

Với các cuộc gọi đã được đánh giá, hệ thống có thể đưa ra các phân tích tổng quan thông qua việc trực quan hóa dữ liệu (Data Visualization). Hệ thống sẽ hiển thị các biểu đồ và các con số thống kê mang tính quan trọng cho các nhà quản lý, ví dụ như tỷ lệ khách hàng hài lòng, số lượng cuộc gọi nhân viên đạt yêu cầu, số lượng cuộc gọi đã được đánh giá...

Hệ thống cũng đưa ra các phân tích chi tiết của từng nhân viên CSKH, từ đó chỉ ra các điểm mạnh và điểm yếu của từng nhân viên để quản lý viên có thể đánh giá.

3.4.3 Tính năng số 3: Đào tạo nghiệp vụ chăm sóc khách hàng.

Hệ thống cho phép đào tạo nghiệp vụ chăm sóc khách hàng cho các nhân viên chăm sóc khách hàng cũng như các bạn sinh viên có nhu cầu tăng cường kinh nghiệm bản thân. Trong quá trình đào tạo, người học được yêu cầu nói lại các câu nói theo kịch bản mà hệ thống đưa ra. Các câu nói trong kịch bản được xây dựng theo chủ đề nghiệp vụ chăm sóc khách hàng.

Sau khi câu nói của người học đã được ghi âm lại trên hệ thống, hệ thống sẽ tiến hành đánh giá so khớp nội dung của câu nói đó và câu nói trong kịch bản, cũng như nhận diện cảm xúc trong câu nói. Nếu như nội dung được cung cấp khớp với câu nói trong kịch bản, và cảm xúc trong câu nói mang tính tích cực, hệ thống sẽ đánh giá đạt cho câu nói của người học và chuyển sang các câu nói khác trong kịch bản. Nếu như câu nói của người học không đạt, người học sẽ được yêu cầu nói lại câu nói đó trước khi chuyển sang câu tiếp theo.

3.5 Phân tích chuyên sâu sản phẩm.

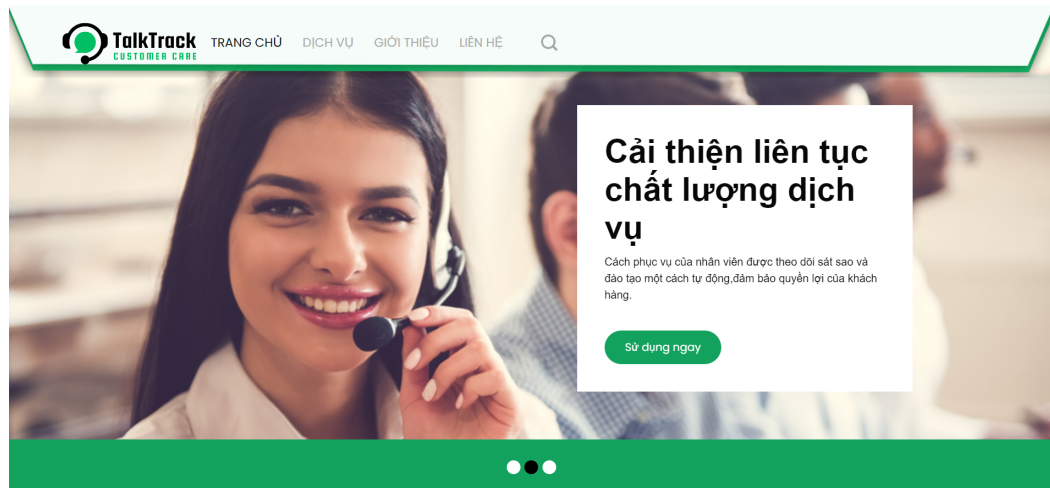
3.5.1 Tính đổi mới, tối ưu trong công nghệ.

- Sản phẩm TalkTrack hỗ trợ tự động hóa quy trình đào tạo nghiệp vụ chăm sóc khách hàng cho nhân viên.
- Các mô hình AI được huấn luyện tập trung theo dữ liệu tiếng nói tiếng Việt nói chung và theo chủ đề chăm sóc khách hàng nói riêng, từ đó đảm bảo độ chính xác trong quá trình vận hành.
- TalkTrack có thể giúp nhà quản lý dễ dàng nắm bắt nội dung cuộc gọi thông qua các từ khóa được trích xuất mà không cần phải nghe hay xem hết nội dung văn bản của cuộc gọi.
- Việc chấm điểm đánh giá cho nhân viên dựa trên nhiều tiêu chí khác nhau để đảm bảo tính khách quan, ví dụ như số lượng khách hàng mà hài lòng với nhân viên đó, hay điểm thái độ trung bình của nhân viên trong tháng.
- Toàn bộ hệ thống được thiết kế để bảo mật dữ liệu về doanh nghiệp, người dùng, đảm bảo dữ liệu của khách hàng được sử dụng đúng mục đích, không cung cấp dữ liệu cho bên thứ ba.

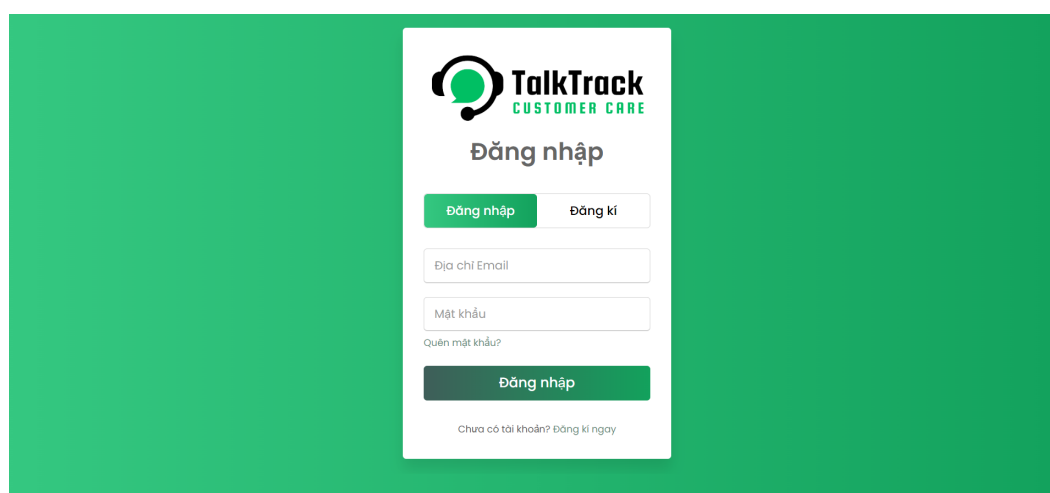
3.5.2 Trải nghiệm người dùng.

Dưới đây là các mô tả chi tiết về trải nghiệm người dùng trong bản Demo của sản phẩm TalkTrack. Ở thời điểm hiện tại, người dùng có thể truy cập và dùng thử bản Demo tại đường dẫn sau đây: <https://thanhpv2102.github.io/sitas/>

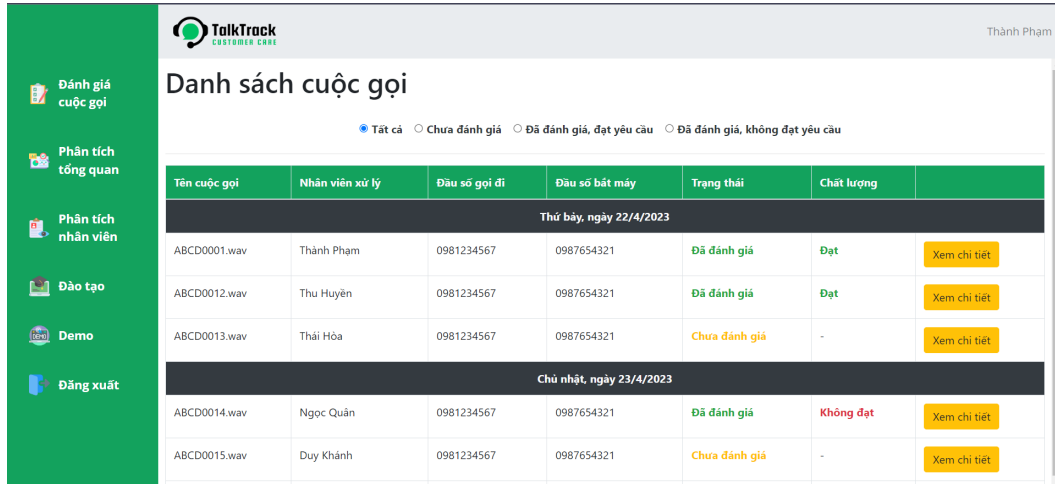
Hình 3.1: Màn hình trang chủ của TalkTrack. Trang chủ sẽ hiển thị thông tin chi tiết về các tính năng và dịch vụ của TalkTrack.



Hình 3.2: Màn hình trang đăng nhập / đăng ký. Sau khi bấm vào nút 'Sử dụng ngay' ở trang chủ, website sẽ trở đến trang đăng nhập / đăng ký như hình mô tả dưới đây:

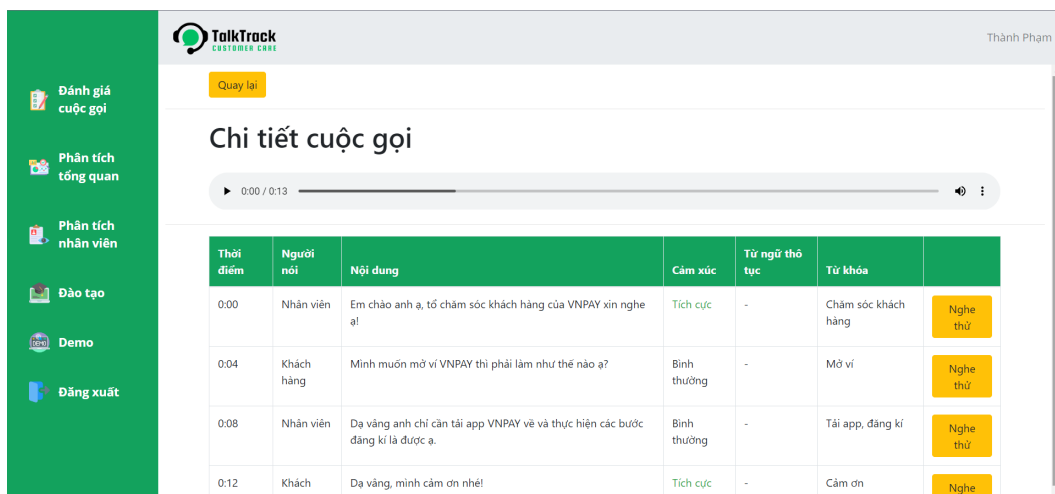


Hình 3.3: Màn hình trang đánh giá chất lượng cuộc gọi. Sau khi bấm nút 'Đăng nhập', màn hình trang đánh giá được hiển thị. Tại đây, người dùng có thể xem được danh sách các cuộc gọi trên hệ thống. Ngoài ra, sẽ có một thanh điều hướng nằm dọc phía bên trái màn hình để người dùng có thể chọn sử dụng các tính năng khác của hệ thống.



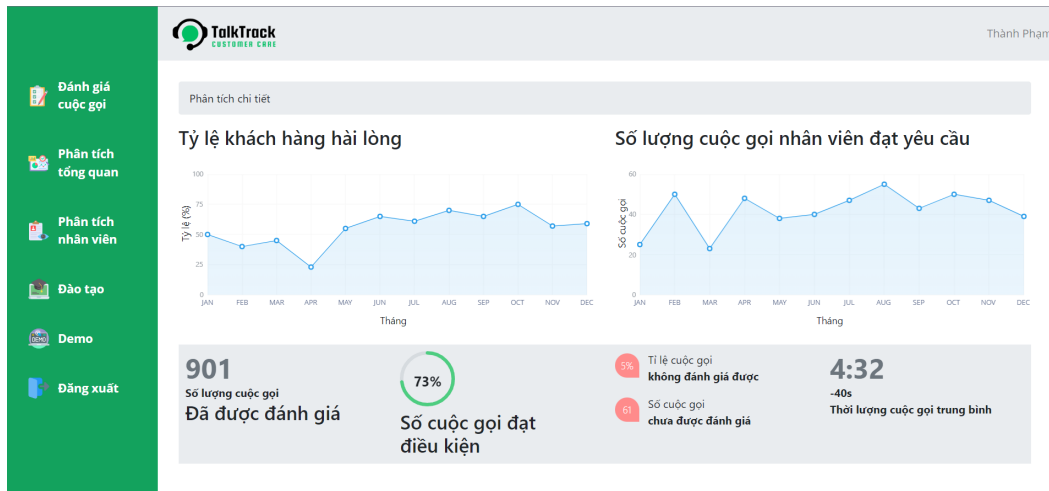
Tên cuộc gọi	Nhân viên xử lý	Đầu số gọi đi	Đầu số bắt máy	Trạng thái	Chất lượng	
Thứ bảy, ngày 22/4/2023						
ABCD0001.wav	Thành Phạm	0981234567	0987654321	Đã đánh giá	Đạt	Xem chi tiết
ABCD0012.wav	Thu Huyền	0981234567	0987654321	Đã đánh giá	Đạt	Xem chi tiết
ABCD0013.wav	Thái Hòa	0981234567	0987654321	Chưa đánh giá	-	Xem chi tiết
Chủ nhật, ngày 23/4/2023						
ABCD0014.wav	Ngọc Quân	0981234567	0987654321	Đã đánh giá	Không đạt	Xem chi tiết
ABCD0015.wav	Duy Khánh	0981234567	0987654321	Chưa đánh giá	-	Xem chi tiết

Hình 3.4: Màn hình trang chi tiết cuộc gọi. Trang này sẽ hiển thị nội dung của từng câu nói trong cuộc đối thoại, bao gồm nội dung văn bản, người nói, từ khóa, từ ngữ thô tục và cảm xúc. Ở đây người dùng cũng có thể nghe thử audio để đánh giá.

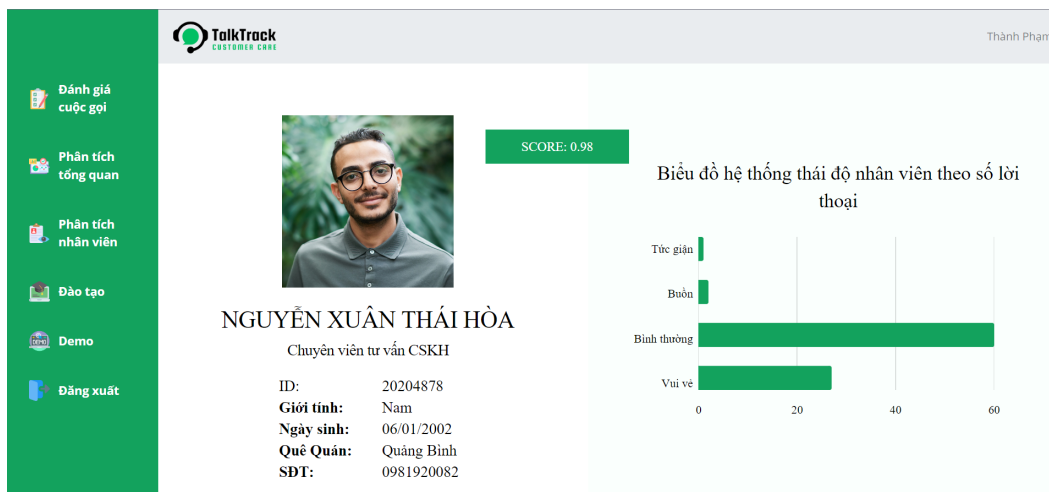


Thời điểm	Người nói	Nội dung	Cảm xúc	Từ ngữ thô tục	Từ khóa	
0:00	Nhân viên	Em chào anh ạ, tổ chăm sóc khách hàng của VNPAY xin nghe ạ!	Tích cực	-	Chăm sóc khách hàng	Nghe thử
0:04	Khách hàng	Mình muốn mở ví VNPAY thì phải làm như thế nào ạ?	Bình thường	-	Mở ví	Nghe thử
0:08	Nhân viên	Dạ vâng anh chỉ cần tải app VNPAY về và thực hiện các bước đăng kí là được ạ.	Bình thường	-	Tải app, đăng kí	Nghe thử
0:12	Khách hàng	Dạ vâng, mình cảm ơn nhé!	Tích cực	-	Cảm ơn	Nghe

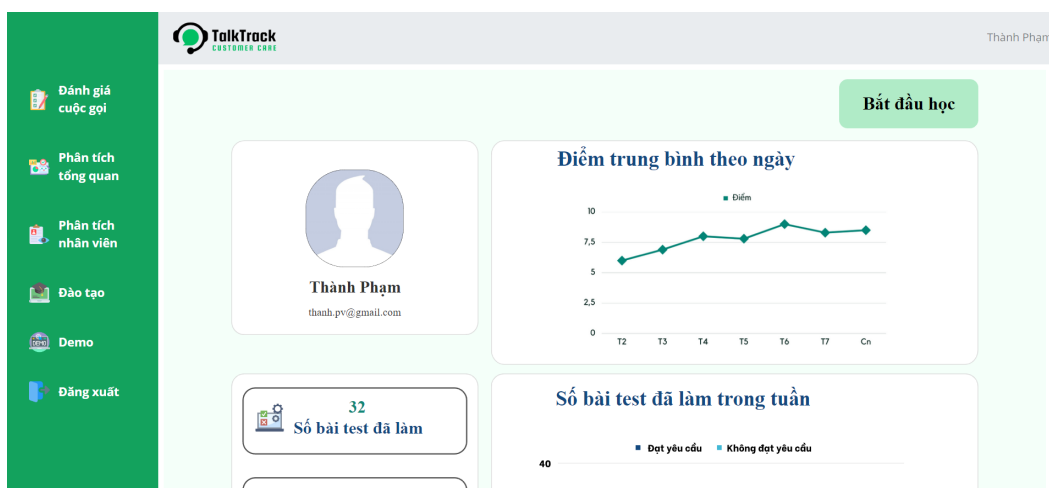
Hình 3.5: Màn hình trang phân tích tổng quan. Ở đây, hệ thống sẽ hiển thị các biểu đồ và các con số thống kê mang tính quan trọng cho các quản lý viên.



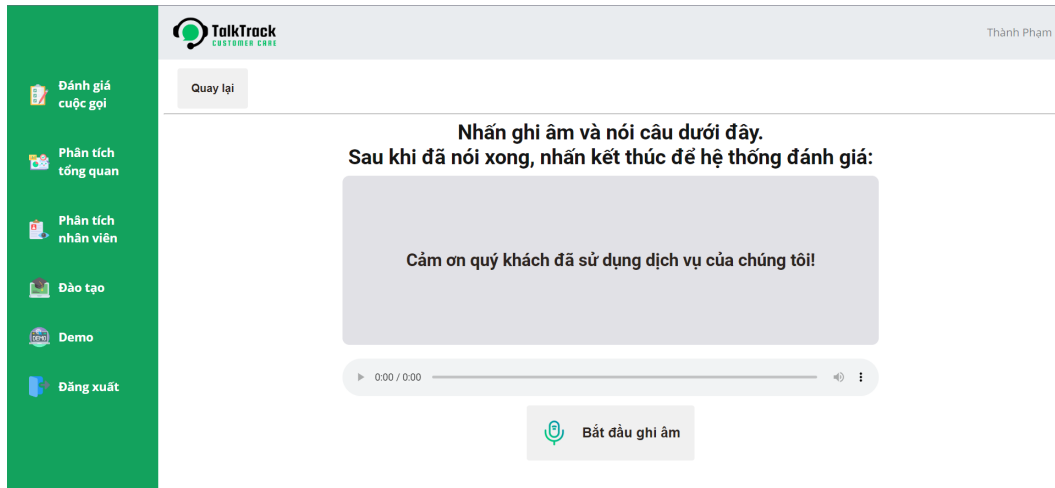
Hình 3.6: Màn hình trang phân tích nhân viên. Trang này sẽ hiển thị một số thông tin cơ bản của nhân viên, cũng như thống kê chất lượng nghiệp vụ hiện tại của nhân viên.



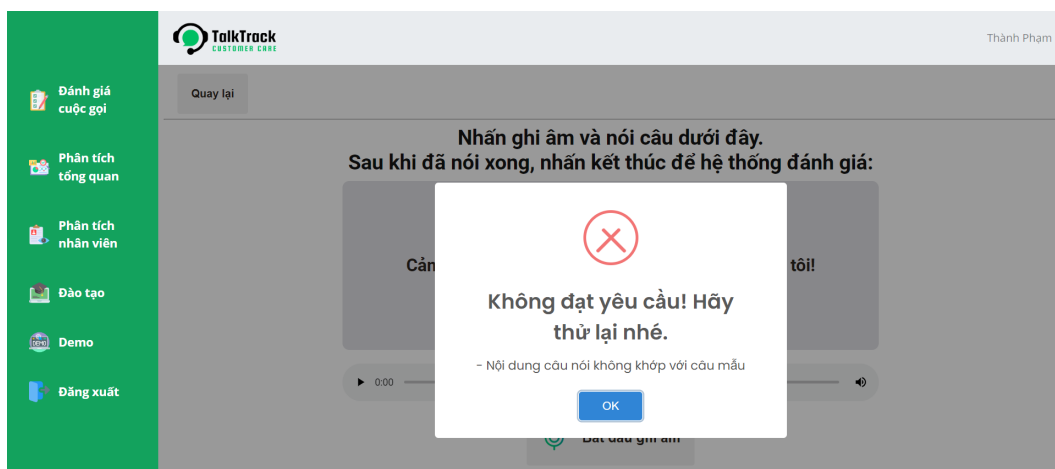
Hình 3.7: Màn hình trang đào tạo. Ở màn hình này, hệ thống sẽ hiển thị các thống kê như số bài test học viên đã làm hay điểm số trung bình.



Hình 3.8: Màn hình trang tự học. Ở màn hình này, các học viên có thể tự tra dồi nghiệp vụ chăm sóc khách hàng của bản thân. Ở đây, hệ thống sẽ hiển thị 1 câu nói có trong kịch bản và yêu cầu học viên nói lại câu nói đó.



Hình 3.9: Màn hình trang đào tạo khi học viên chưa đạt yêu cầu về nội dung. Khi câu nói của học viên không khớp với nội dung cho trước, hệ thống sẽ yêu cầu học viên nói lại câu nói đó trước khi chuyển sang câu nói tiếp theo.



3.5.3 Tiềm năng phát triển rộng rãi của sản phẩm.

Với những gì mà TalkTrack có thể làm được ở thời điểm hiện tại, trong tương lai, nó hoàn toàn có thể được cải tiến và mở rộng để có thể thực hiện được nhiều chức năng hơn và tiền gần hơn tới nhiều doanh nghiệp khác hơn nữa. Cụ thể:

- Chúng ta có thể bổ sung thêm tập dữ liệu Tiếng Anh để huấn luyện mô hình để TalkTrack có thể dùng cho dịch vụ các cuộc đối thoại quốc tế, bắt kịp với xu hướng đối thoại xen lẫn Tiếng Anh và Tiếng Việt.

- Thêm tính năng xử lý trong thời gian thực: các cuộc gọi sẽ được giám sát và xử lý ngay khi đang được thực hiện thay vì chờ đến khi cuộc gọi kết thúc.

- Kết hợp đồng thời các mô hình AI khác như phân loại ý định khách hàng,... nhằm tăng tính ứng dụng của sản phẩm.

- Bổ sung thêm vào tập dữ liệu huấn luyện nhiều hơn nữa, đặc biệt là các dữ liệu chuyên sâu trong ngành dịch vụ, để mô hình có thể học tốt hơn và đưa ra kết quả chính xác hơn.

Trong tương lai, team sẽ thu thập thêm các đoạn ghi âm trực tiếp hội thoại từ các doanh nghiệp khác, đặc biệt là các doanh nghiệp nước ngoài và thực hiện gán nhãn để tạo ra tập dữ liệu bổ sung chính xác hơn. Đồng thời áp dụng các mô hình tiên tiến cho bài toán phân loại ý định khách hàng để lồng ghép vào các bài toán hiện tại.

Chương 4

Nguồn lực

4.1 Đội ngũ nhân sự.

Phạm Việt Thành:

- Kinh nghiệm 3 năm nghiên cứu về lĩnh vực NLP.
- Đạt nhiều giải thưởng và nghiên cứu khoa học về lĩnh vực xử lý tiếng nói.
- Kinh nghiệm 2 năm làm AI Engineer ở công ty start-up về chuyển đổi giọng nói thành văn bản.
- IELTS 7.5

Ngô Thị Thu Huyền:

- Kinh nghiệm 1 năm học tập, nghiên cứu về lĩnh vực NLP.
- Đạt học bổng khuyến khích học tập đại học Bách Khoa.
- Top 3 cuộc thi O-COCOSDA and VLSP 2022 Challenge on Multilingual Speaker Verification.

Nguyễn Xuân Thái Hòa:

- Kinh nghiệm 1 năm học tập, nghiên cứu về lĩnh vực NLP.
- Giải nhì tiếng anh cấp quốc gia.
- Đạt học bổng khuyến khích học tập đại học Bách Khoa.
- Top 2 cuộc thi O-COCOSDA and VLSP 2022 Challenge on Multilingual Speaker Verification.

Phạm Ngọc Quân:

- Top 1 cuộc thi sinh viên SV Start Up.
- Giải ba học sinh giỏi quốc gia môn toán .

Nguyễn Duy Khánh:

- Kinh nghiệm 1 năm học tập, nghiên cứu về lĩnh vực NLP,
- Công bố khoa học ở hội nghị khoa học thường niên về lĩnh vực xử lý âm thanh.
- Top 2 cuộc thi O-COCOSDA and VLSP 2022 Challenge on Multilingual Speaker Verification.

4.2 Nguồn lực nhân sự khác.

Đơn vị hỗ trợ: Lab NLP, Đại học Bách khoa Hà Nội.

Cố vấn chiến lược: Mai Linh - sinh viên năm 3 ngành quản trị kinh doanh, trường đại học FPT.

4.3 Nguồn lực tài chính và cơ sở hạ tầng khác.

Cung cấp cơ sở hạ tầng : Server cloud cung cấp bởi lab NLP, Đại học Bách khoa Hà Nội.