

**BÁO CÁO NGHIÊN CỨU**

**Tên dự án: Musem**

**—o0o—**

**Học sinh 1. Lê Phú Trọng SBD: BDAI-011**

**Học sinh 2. Phạm Văn Minh Phúc SBD: BDAI-021**

**Học sinh 3. Vũ Gia Bảo SBD: BDAI-045**

**Học sinh 4. Nguyễn Vũ An Vương SBD: BDAI-055**

**Học sinh 5. Nguyễn Phạm Đức Huy SBD: BDAI-056**

**Thành phố Hồ Chí Minh, tháng 9 năm 2025**

---

# Contents

<b>1</b>	<b>Mở đầu</b>	<b>1</b>
1.1	Đặt vấn đề . . . . .	1
1.2	Các giải pháp hiện tại . . . . .	1
1.3	Giải pháp của nhóm . . . . .	1
<b>2</b>	<b>Tiền xử lí dữ liệu</b>	<b>3</b>
2.1	Nguồn thu thập . . . . .	3
2.2	Đọc và tiền xử lí . . . . .	3
2.3	Trích xuất đặc trưng . . . . .	3
2.4	Dán nhãn . . . . .	4
<b>3</b>	<b>Huấn luyện mô hình</b>	<b>5</b>
3.1	Pipeline huấn luyện . . . . .	5
3.2	Tối ưu siêu tham số . . . . .	5
3.3	Mã nguồn minh hoạ . . . . .	5
3.4	Đánh giá . . . . .	6
3.4.1	Kết quả chi tiết . . . . .	6
3.4.2	Ma trận nhầm lẫn . . . . .	6
3.4.3	Phân tích kết quả . . . . .	6
3.4.4	Kết luận . . . . .	7
<b>4</b>	<b>Xây dựng phần mềm</b>	<b>8</b>
4.1	Giới thiệu hệ thống . . . . .	8
4.2	Mô hình . . . . .	8
4.3	Dữ liệu . . . . .	8
4.4	Xây dựng giao diện . . . . .	8
4.5	Xử lý giọng nói . . . . .	8
<b>5</b>	<b>Vấn đề đạo đức và bảo mật trong AI tư vấn tâm lí</b>	<b>9</b>
5.1	Bảo mật dữ liệu . . . . .	9
5.2	Đạo đức . . . . .	9
<b>6</b>	<b>Tổng kết</b>	<b>10</b>
6.1	Hạn chế . . . . .	10
6.2	Hướng phát triển . . . . .	10
6.3	Ý nghĩa . . . . .	10
<b>7</b>	<b>Tài liệu tham khảo</b>	<b>11</b>

---

# 1 Mở đầu

## 1.1 Đặt vấn đề

Trong kỷ nguyên số hiện nay, âm nhạc đã trở thành một phần không thể thiếu trong đời sống tinh thần của nhiều người. Âm nhạc không chỉ là một hình thức thư giãn, giải trí mà người nghe nhạc còn tìm được sự đồng điệu về mặt cảm xúc trong các bản nhạc, từ đó giải tỏa tâm lý và phục hồi tinh thần, cực kì hữu dụng trong cuộc sống hiện đại đầy xô bồ và áp lực. Không những thế, các nghiên cứu trong lĩnh vực Tâm lý học cũng chỉ ra rằng con người có xu hướng tìm đến những bản nhạc phù hợp với tâm trạng của mình, hoặc đôi khi là những giai điệu có khả năng chuyển đổi cảm xúc theo hướng tích cực hơn.

Tuy nhiên, việc tự mình tìm kiếm các bản nhạc phù hợp giữa một kho tàng âm nhạc khổng lồ trên các nền tảng nghe nhạc trực tuyến lại trở thành một thách thức không nhỏ về thời gian và công sức, đặc biệt là trong những khoảnh khắc người dùng đang trải qua các cung bậc cảm xúc mạnh mẽ.

## 1.2 Các giải pháp hiện tại

Các nền tảng âm nhạc trực tuyến hiện nay đã triển khai hệ thống gợi ý nhạc dựa trên thói quen nghe nhạc của người dùng hay theo ngữ cảnh (như thời gian, địa điểm). Tuy nhiên, các hệ thống này chỉ dựa trên dữ liệu lịch sử hay yếu tố ngoại cảnh chứ chưa khai thác yếu tố cảm xúc của người dùng tại thời điểm nghe nhạc. Ngoài ra, các hệ thống này có xu hướng gợi ý liên tục một kiểu nhạc giống nhau, hạn chế sự đa dạng và vô tình "nhốt" người dùng trong một trạng thái cảm xúc tiêu cực thay vì giúp họ chuyển đổi sang trạng thái tích cực hơn.

Một số ứng dụng cho phép người dùng tự chọn biểu tượng cảm xúc (emoji) hoặc gõ từ khoá để hệ thống trả về các playlist nhạc tương ứng. Dẫu vậy, việc buộc người dùng phải tự mô tả cảm xúc hiện tại của mình là một rào cản lớn, làm gián đoạn trải nghiệm và cảm xúc được khai báo có thể không chính xác do người dùng không muốn thừa nhận hay không biết diễn tả cảm xúc như thế nào.

## 1.3 Giải pháp của nhóm

Các giải pháp nêu trên thiếu đi một cơ chế thu thập và phân tích tín hiệu cảm xúc thời gian thực một cách thụ động và chân thực. Sự vắng mặt của một hệ thống có khả năng "lắng nghe" và "thấu hiểu" trạng thái cảm xúc thông qua các tín hiệu sinh học tự nhiên đã được khoa học chứng minh là phản ánh cảm xúc con người với độ chính xác cao (như giọng nói) chính là khoảng trống lớn cần được lấp đầy.

Chính vì vậy, nhóm xây dựng một hệ thống chatbot gọi là **Musem (Music - Emotion)**. Hệ thống cung cấp khả năng đề xuất nhạc theo cảm xúc của người dùng dựa trên các đặc trưng âm thanh của **giọng nói**.

Hệ thống gồm 3 phần:

- Một chatbot có khả năng trò chuyện với người dùng bằng giọng nói cũng như tin nhắn văn bản. Người dùng trò chuyện với chatbot này bằng giọng nói.
- Một mô hình AI có khả năng phân tích cảm xúc hiện tại của người dùng thông qua các đặc trưng âm thanh trong giọng nói.

- 
- Một hệ thống gợi ý nhạc cho người dùng bằng thuật toán truyền thống dựa trên kết quả phân tích cảm xúc.

Hệ thống hứa hẹn mang lại một giải pháp tiện lợi, liền mạch và thực sự thấu hiểu người dùng.

---

## 2 Tiền xử lí dữ liệu

### 2.1 Nguồn thu thập

Các bộ dữ liệu có sẵn trên nền tảng Kaggle:

- Để phân tích cảm xúc, ta sẽ sử dụng bộ dữ liệu [RAVDESS Emotional Speech Audio](#). Bộ dữ liệu này bao gồm 1440 tập tin âm thanh có định dạng .wav, với 24 diễn viên (12 nam, 12 nữ). Mỗi diễn viên sẽ đọc 2 câu tiếng Anh với nhiều cảm xúc khác nhau như: bình thường, vui, buồn, tức giận, sợ hãi, bất ngờ, ghê tởm.
- Để gợi ý bài nhạc, ta sẽ sử dụng bộ dữ liệu [Spotify Music Data to Identify the Moods](#). Bộ dữ liệu này chứa các thông tin thuộc tính của bài hát (ví dụ: năng lượng, tempo, valence, danceability, loudness) được gắn nhãn theo tâm trạng, giúp mô hình có thể đưa ra gợi ý bài hát phù hợp với cảm xúc người dùng.

### 2.2 Đọc và tiền xử lí

Để xử lí tín hiệu âm thanh, ta sử dụng thư viện librosa (McFee et al., 2015). Quá trình đọc một tập tin âm thanh và tiền xử lí cơ bản bao gồm các bước sau:

```
1 import librosa
2 y, sr = librosa.load("audio.wav", sr=None)
```

Trong đó:

- **y**: là một mảng NumPy chứa dữ liệu sóng âm (biên độ theo thời gian), đã được chuẩn hóa vào khoảng  $[-1, 1]$ .
- **sr** (sampling rate): tần số lấy mẫu, đơn vị là Hz (số mẫu/giây). Giá trị này cho biết độ phân giải thời gian của tín hiệu. Mặc định, librosa.load sẽ đưa về 22050 Hz, nhưng với tham số sr=None ta có thể giữ nguyên tần số gốc của file.

Để cải thiện chất lượng dữ liệu đầu vào, một số bước tiền xử lí cơ bản được áp dụng:

- **Chuẩn hóa**: đảm bảo các giá trị sóng âm nằm trong khoảng  $[-1, 1]$  để ổn định quá trình huấn luyện.
- **Cắt bỏ khoảng lặng (trimming)**: loại bỏ các đoạn im lặng ở đầu và cuối file nhằm tập trung vào phần tín hiệu hữu ích. Ví dụ:

```
1 y_trimmed, index = librosa.effects.trim(y)
```

- **Cân bằng độ dài**: đưa tất cả các mẫu âm thanh về cùng một độ dài (padding hoặc cắt bớt), giúp quá trình trích xuất đặc trưng thống nhất.

Kết quả của bước này là một tập dữ liệu sóng âm đã được chuẩn hóa, làm sạch và sẵn sàng cho quá trình trích xuất đặc trưng.

### 2.3 Trích xuất đặc trưng

Các tập tin âm thanh trong bộ dữ liệu RAVDESS ở dạng sóng âm .wav. Để có thể đưa dữ liệu vào mô hình học máy (SVM), cần chuyển đổi tín hiệu âm thanh thành đặc trưng số.

---

Trong đề tài này, ta sử dụng **MFCC (Mel-Frequency Cepstral Coefficients)** làm đặc trưng chính (Davis, S. and Mermelstein, P., 1980). MFCC là một phương pháp phổ biến trong xử lý giọng nói, dựa trên thang đo tần số Mel nhằm mô phỏng cách tai người cảm nhận âm thanh. Quy trình tính toán MFCC bao gồm các bước:

1. Chia tín hiệu âm thanh thành nhiều khung nhỏ (frames).
2. Thực hiện biến đổi Fourier ngắn hạn (STFT) để đưa tín hiệu sang miền tần số.
3. Áp dụng thang đo Mel để mô phỏng sự nhạy cảm của tai người với các dải tần số khác nhau.
4. Lấy log năng lượng của các dải tần số.
5. Thực hiện biến đổi Cosine rời rạc (DCT) để thu được các hệ số MFCC.

Kết quả cuối cùng của mỗi file âm thanh là một vector MFCC chứa 13 hệ số, với mỗi hệ số ta sẽ lưu giá trị trung bình (mean) và độ lệch chuẩn (std) và sử dụng làm đầu vào cho mô hình SVM nhằm phân loại cảm xúc trong giọng nói.

## 2.4 Dán nhãn

Trong bộ dữ liệu RAVDESS, mỗi tập tin âm thanh đã được gắn nhãn cảm xúc tương ứng. Các nhãn này bao gồm: *neutral*, *calm*, *happy*, *sad*, *angry*, *fearful*, *disgust*, *surprised*. Khi huấn luyện mô hình, nhãn sẽ được ánh xạ về dạng số (ví dụ: 0-7) để phù hợp với quá trình phân loại.

Đối với dữ liệu Spotify, mỗi bài hát đã được gắn nhãn mood hoặc mức độ valence (từ tiêu cực đến tích cực). Các nhãn này sẽ đóng vai trò làm cơ sở cho việc gợi ý bài nhạc phù hợp với cảm xúc được nhận diện từ giọng nói của người dùng.

---

## 3 Huấn luyện mô hình

Trong phần này, chúng tôi trình bày chi tiết quá trình huấn luyện mô hình nhận diện cảm xúc giọng nói (Speech Emotion Recognition – SER) dựa trên phương pháp **Support Vector Machine (SVM)** với đặc trưng được trích xuất bằng thư viện **Librosa**. Mục tiêu là xây dựng bộ phân loại có khả năng tổng quát hóa tốt trên các diễn giả khác nhau.

### 3.1 Pipeline huấn luyện

Toàn bộ quy trình được đóng gói trong pipeline để tránh rò rỉ thông tin:

1. Chuẩn hoá dữ liệu bằng StandardScaler.
2. (Tuỳ chọn) Giảm chiều bằng PCA, giữ lại 95
3. Huấn luyện bộ phân loại SVM với kernel RBF. Tham số  $C$  và  $\gamma$  được hiệu chỉnh qua tìm kiếm có lưới (grid search). Tham số `class_weight=balanced` để xử lý mất cân bằng dữ liệu.

### 3.2 Tối ưu siêu tham số

Tiến hành tìm kiếm trên lưới:

$$C \in 10^{-2}, 10^{-1}, 1, 10, 10^2,$$
$$\gamma \in 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1.$$

Tiêu chí chọn mô hình là macro-F1 để cân bằng giữa các lớp. Kết quả tốt nhất trong cross-validation sẽ được huấn luyện lại trên toàn bộ tập train và đánh giá trên tập test.

### 3.3 Mã nguồn minh hoạ

```
1 from sklearn.svm import SVC
2 from sklearn.pipeline import Pipeline
3 from sklearn.preprocessing import StandardScaler
4 from sklearn.decomposition import PCA
5 from sklearn.model_selection import GridSearchCV, StratifiedKFold
6 from sklearn.model_selection import train_test_split
7 from sklearn.metrics import classification_report, confusion_matrix,
  ConfusionMatrixDisplay
8
9 pipeline = Pipeline([('scaler', StandardScaler()),
10                      ('pca', PCA(n_components=0.95)),
11                      ('svm', SVC(kernel='rbf', class_weight='balanced'))])
12 param_grid = {'svm__C': [0.01, 0.1, 1, 10, 100],
13               'svm__gamma': [1e-4, 1e-3, 1e-2, 1e-1, 1]}
14
15 cv = StratifiedKFold(n_splits=5, shuffle=True, random_state=42)
16 model = GridSearchCV(pipeline, param_grid, scoring='f1_macro', cv=cv)
17 model.fit(X_train, Y_train)
```

### 3.4 Đánh giá

Mô hình SVM được huấn luyện trên tập dữ liệu giọng nói gán nhãn cảm xúc, sau đó đánh giá trên tập kiểm thử độc lập. Các chỉ số **Precision**, **Recall**, **F1-score** và **Accuracy** được sử dụng để phản ánh hiệu năng của mô hình trên từng lớp cảm xúc cũng như toàn bộ hệ thống.

#### 3.4.1 Kết quả chi tiết

Bảng 1 trình bày kết quả phân loại chi tiết theo từng lớp cảm xúc.

**Table 1:** Báo cáo phân loại của mô hình SVM trên tập kiểm thử

Cảm xúc	Precision	Recall	F1-score	Số mẫu
Neutral	0.52	0.63	0.57	19
Calm	0.69	0.87	0.77	38
Happy	0.81	0.68	0.74	38
Sad	0.70	0.61	0.65	38
Angry	0.86	0.77	0.81	39
Fearful	0.67	0.77	0.71	39
Disgust	0.75	0.63	0.69	38
Surprised	0.65	0.67	0.66	39
<b>Accuracy</b>		0.71 (288 mẫu)		
<b>Macro Avg</b>	0.71	0.70	0.70	–
<b>Weighted Avg</b>	0.72	0.71	0.71	–

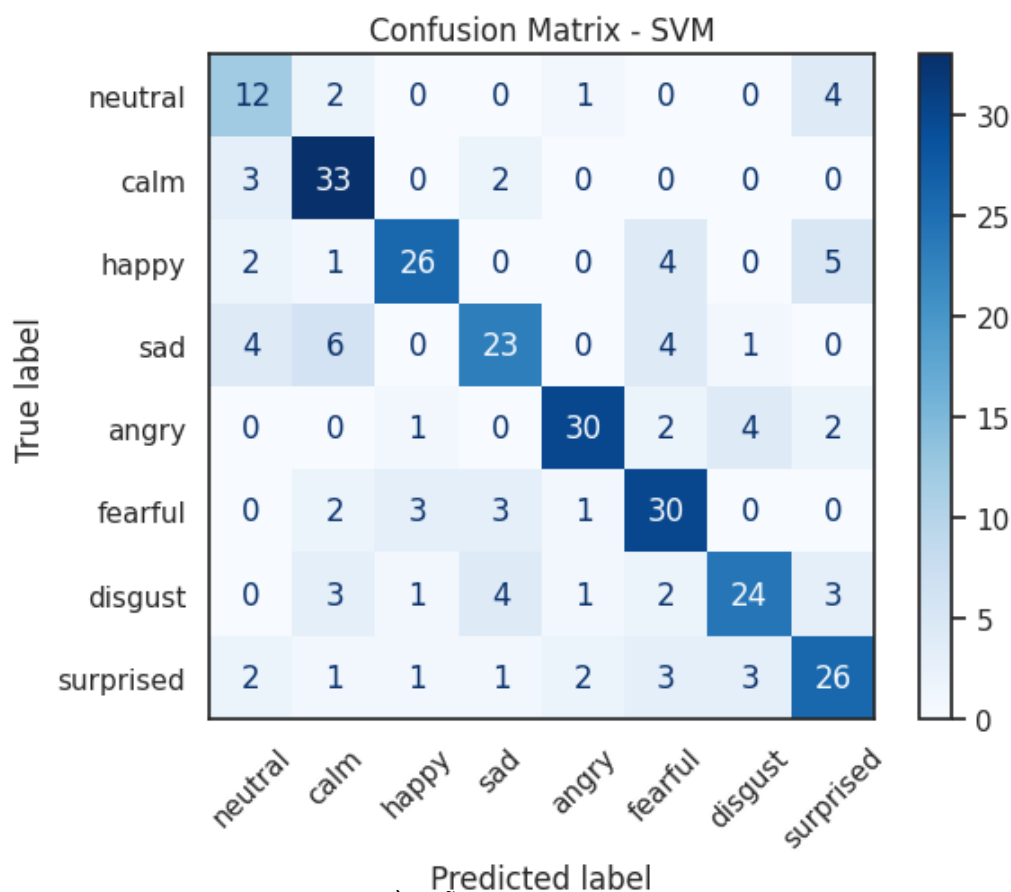
#### 3.4.2 Ma trận nhầm lẫn

Để trực quan hơn về các lỗi phân loại, Hình 1 minh họa ma trận nhầm lẫn (confusion matrix). Ta thấy rằng các cảm xúc như *calm*, *angry* và *happy* được phân loại khá tốt, trong khi *neutral* và *sad* có tỷ lệ nhầm lẫn cao hơn.

#### 3.4.3 Phân tích kết quả

- Mô hình đạt **Accuracy 71%**, với **Macro-F1 = 0.70**, cho thấy hiệu năng khá cân bằng giữa các lớp cảm xúc.
- Các cảm xúc có đặc trưng rõ rệt như *angry* (F1 = 0.81) và *happy* (F1 = 0.74) được phân loại tốt, với precision cao.
- Cảm xúc *calm* đạt Recall tới 0.87, nghĩa là hầu hết mẫu *calm* được nhận diện chính xác.
- Ngược lại, *neutral* (F1 = 0.57) và *sad* (F1 = 0.65) cho thấy mức độ nhầm lẫn cao hơn, do đặc trưng âm học gần với các cảm xúc trung tính khác.
- Cảm xúc *fearful* và *surprised* có recall khá cao (0.77 và 0.67), chứng tỏ mô hình ít bỏ sót nhưng precision chưa thật sự tối ưu.





**Figure 1:** Ma trận nhầm lẫn của mô hình SVM cho bài toán SER

#### 3.4.4 Kết luận

Mô hình SVM cho SER đã đạt hiệu năng tốt trên tập kiểm thử, với độ chính xác tổng thể trên 70%. Điểm mạnh là khả năng nhận diện các cảm xúc mạnh (angry, happy, calm), tuy nhiên vẫn tồn tại sự nhầm lẫn ở các cảm xúc gần nhau về đặc trưng âm học (neutral, sad, disgust).

---

## 4 Xây dựng phần mềm

### 4.1 Giới thiệu hệ thống

Đề tài của nhóm tập trung xây dựng một hệ thống chatbot có khả năng gợi ý bài hát dựa trên cảm xúc người dùng. Xuất phát từ việc âm nhạc có tác động mạnh mẽ đến cảm xúc của con người. Khi người dùng tương tác với chatbot, hệ thống sẽ phân tích tâm trạng của người dùng rồi đưa ra lời khuyên kèm gợi ý bài hát để người dùng có thể trở nên tốt hơn.

### 4.2 Mô hình

Hệ thống được phát triển dựa trên mô hình **Gemini AI** kết hợp với phương pháp **Retrieval-Augmented Generation (RAG)**. Mô hình đảm nhiệm vai trò phân tích và phản hồi theo ngữ cảnh, đồng thời được tối ưu để nhận diện cảm xúc trong âm nhạc và hội thoại.

### 4.3 Dữ liệu

Để huấn luyện và tinh chỉnh mô hình, nhóm sử dụng bộ dữ liệu từ quyển sách tâm lý và học đường:

- <https://www.scribd.com/document/828336318/Sach-Tam-ly-hoc-duong>

Cuốn sách tập trung vào tâm lý học đường và thực hành tư vấn trong môi trường giáo dục, nhằm hỗ trợ sự phát triển toàn diện của học sinh về mặt cá nhân và xã hội

### 4.4 Xây dựng giao diện

Phần giao diện người dùng được xây dựng bằng thư viện **Gradio**. Người dùng có thể tương tác trực tiếp với hệ thống thông qua trình duyệt web, bao gồm các chức năng:

- Ghi âm và gửi giọng nói cho chatbot phân tích.
- Nhận phản hồi từ AI dưới dạng văn bản và âm thanh.
- Trải nghiệm hệ thống một cách trực quan và thân thiện.

### 4.5 Xử lý giọng nói

Hệ thống tích hợp các thư viện xử lý giọng nói để tăng tính tự nhiên trong tương tác:

- **SpeechRecognition**: Dùng để chuyển đổi giọng nói người dùng thành văn bản phục vụ mô hình phân tích.
- **gTTS (Google Text-to-Speech)**: Dùng để sinh giọng nói phản hồi từ mô hình, tạo trải nghiệm giao tiếp hai chiều.

---

## 5 Vấn đề đạo đức và bảo mật trong AI tư vấn tâm lý

### 5.1 Bảo mật dữ liệu

Trong các hệ thống tư vấn tâm lý dựa trên AI, dữ liệu người dùng có tính chất đặc biệt nhạy cảm, bao gồm cảm xúc, tình trạng tâm lý, thông tin cá nhân và gia đình. Do đó, việc bảo mật dữ liệu là yêu cầu then chốt. Các biện pháp cần được thực hiện bao gồm:

- **Ẩn danh (Anonymization):** Loại bỏ thông tin có thể định danh nhằm bảo vệ quyền riêng tư người dùng.
- **Giảm thiểu dữ liệu thu thập:** Chỉ lưu trữ những thông tin cần thiết phục vụ mô hình, tránh thu thập dư thừa.
- **Tuân thủ pháp lý:** Các hệ thống cần tuân thủ các quy định về bảo vệ dữ liệu như GDPR, HIPAA hoặc quy chuẩn pháp luật tại địa phương.

### 5.2 Đạo đức

Bên cạnh bảo mật, vấn đề đạo đức cũng có vai trò quan trọng trong việc ứng dụng AI cho tư vấn tâm lý:

- **Minh bạch và trách nhiệm:** Người dùng cần được biết rõ rằng họ đang tương tác với một hệ thống AI chứ không phải con người. AI chỉ nên đóng vai trò hỗ trợ, không thay thế chuyên gia tâm lý.
- **Thiên lệch dữ liệu (Bias):** Mô hình có thể phản ánh các định kiến xã hội nếu dữ liệu huấn luyện không đa dạng. Điều này cần được kiểm tra và hiệu chỉnh thường xuyên.
- **Giới hạn và hậu quả:** AI không đủ năng lực để chẩn đoán hoặc điều trị các rối loạn tâm lý. Trong các trường hợp khẩn cấp (ví dụ khi người dùng có ý định tự tử), hệ thống cần có cơ chế cảnh báo và khuyến nghị tìm đến chuyên gia.
- **Khung pháp lý và đạo đức nghề nghiệp:** Cần xác định rõ ràng trách nhiệm pháp lý nếu người dùng chịu tổn hại khi tin tưởng vào AI.

---

## 6 Tổng kết

### 6.1 Hạn chế

#### Về kĩ thuật:

- Mô hình phân tích cảm xúc chưa đạt độ chính xác cao.
- Sử dụng dữ liệu có sẵn trên internet nên bộ dữ liệu chưa đủ nhiều để huấn luyện mô hình đủ tốt.
- Sử dụng thuật toán truyền thống để gợi ý nhạc nên còn nhiều hạn chế.
- Sử dụng Gradio để deploy mô hình lên website nên lịch sử trò chuyện không được lưu.

#### Về ứng dụng:

- Hệ thống không thực sự có kiến thức về y tế và không thể thay thế việc đi trị liệu khi gặp vấn đề về tâm lí, người dùng cần có hiểu biết về sức khỏe tinh thần và tìm kiếm sự giúp đỡ khi cần thiết.
- Hệ quả của việc một người thực sự gặp vấn đề tâm lí lạm dụng, dựa dẫm vào chatbot và thu mình với mọi người xung quanh là vô cùng khôn lường.

### 6.2 Hướng phát triển

- Huấn luyện mô hình phân tích cảm xúc với các bộ dữ liệu lớn hơn để tăng độ chính xác.
- Ứng dụng xử lí ngôn ngữ tự nhiên (NLP) để giúp mô hình phân tích cảm xúc hiểu được ý nghĩa lời nói của người dùng, từ đó đưa ra kết quả chính xác hơn.
- Tích hợp thêm một mô hình AI gợi ý nhạc thông minh (thay thế thuật toán truyền thống đang sử dụng) cũng như thêm các tùy chỉnh như thể loại nhạc, tiết tấu, giai điệu để các bản nhạc được gợi ý phù hợp với cảm xúc và sở thích của người dùng hơn cũng như giúp người dùng có cảm xúc tiêu cực dần cải thiện tâm trạng.
- Sử dụng Flask để deploy mô hình lên website nhằm lưu được lịch sử trò chuyện và làm đẹp UI/UX.
- Gợi ý thêm các đa phương tiện khác như phim ảnh, sách, trò chơi điện tử hay các hoạt động khác.

### 6.3 Ý nghĩa

Hệ thống mở ra một hướng tiếp cận mới cho các nền tảng giải trí và chăm sóc sức khỏe tinh thần:

- Cung cấp trải nghiệm nghe nhạc được cá nhân hoá theo đúng nhu cầu cảm xúc hiện tại của người dùng.
- Hỗ trợ sức khỏe tinh thần bước đầu thông qua việc giúp người dùng điều chỉnh và quản lý cảm xúc của mình bằng âm nhạc.

---

## 7 Tài liệu tham khảo

- [1] McFee, Brian and Raffel, Colin and Liang, Dawen and Ellis, Daniel P. W. and McVicar, Matt and Battenberg, Eric and Nieto, Oriol (2015), [Audio and Music Signal Analysis in Python](#).
- [2] Davis, S. and Mermelstein, P. (1980), [Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences](#)
- [3] Scherer, K. R. (2003), [Vocal communication of emotion: A review of research paradigms - Speech Communication, 40\(1-2\), 227-256](#)
- [4] Juslin, P. N. and Laukka, P. (2003), [Comparison of Communication of emotions in vocal expression and music performance: different channels, same code? - Psychological bulletin, 129\(5\), 770](#)
- [5] Banse, R. and Scherer, K. R. (1996), [Comparison of Acoustic profiles in vocal emotion expression - Journal of personality and social psychology, 70\(3\), 614](#)