

THÔNG TIN CHUNG CỦA NHÓM

- Link YouTube video của báo cáo (tối đa 5 phút):
<https://youtu.be/m5pdCMDsmPY>
- Link slides (dạng .pdf đặt trên Github của nhóm):
<https://github.com/lyphucthanh1010/CH2205.CH190/blob/master/Th%C3%A0nh%20L%C3%BD%20Ph%C3%BAC%20-%20CS2205.FEB2025.DeCuong.Final%20Report.Template.Slide.pdf>
- *Mỗi thành viên của nhóm điền thông tin vào một dòng theo mẫu bên dưới*
- *Sau đó điền vào Đề cương nghiên cứu (tối đa 5 trang), rồi chọn Turn in*
- *Lớp Cao học, mỗi nhóm một thành viên*

- Họ và Tên: Lý Phúc Thành
- MSSV: 240101076



- Lớp: CS2205.CH190
- Tự đánh giá (điểm tổng kết môn): 9.5/10
- Số buổi vắng: 0
- Số câu hỏi QT cá nhân: 4
- Số câu hỏi QT của cả nhóm: 4
- Link Github:
<https://github.com/lyphucthanh1010/CH2205.CH190>

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI (IN HOA)

NHẬN DIỆN ÂM NHẠC CÓ BẢN QUYỀN SỬ DỤNG CHROMAPRINTS VÀ MÔ HÌNH PHÂN LOẠI RANDOM FOREST

TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)

MUSIC COPYRIGHT RECOGNITION USING CHROMAPRINTS AND RANDOM FOREST CLASSIFIER

TÓM TẮT

Dự án nghiên cứu này tập trung vào việc ứng dụng các phương pháp học máy để dự đoán bản quyền âm nhạc thông qua việc sử dụng kỹ thuật Chromaprints và thuật toán Random Forest. Chromaprints, một dạng mã vân tay âm thanh, được sử dụng để tạo ra các đặc trưng đại diện cho bản nhạc. Mã vân tay này giúp nhận diện các đặc điểm âm thanh độc đáo của từng bản nhạc, cho phép phân loại và đối chiếu các bài hát với cơ sở dữ liệu âm nhạc hiện có [3]. Các đặc trưng này sau đó được đưa vào mô hình học máy Random Forest, một thuật toán học sâu mạnh mẽ, để dự đoán liệu một bài hát có đang vi phạm bản quyền hay không. Mô hình Random Forest, với khả năng xử lý tốt các tập dữ liệu lớn và phức tạp, giúp tăng cường độ chính xác trong việc phân loại các bài hát dựa trên các đặc trưng âm thanh được trích xuất từ Chromaprints. Quá trình này bao gồm việc thu thập dữ liệu âm nhạc từ nhiều nguồn khác nhau, trích xuất mã vân tay âm thanh, sau đó huấn luyện mô hình học máy để phân biệt giữa các bài hát có bản quyền và không có bản quyền. Mục tiêu của dự án là cung cấp một giải pháp hiệu quả và tự động trong việc xác định và phân loại âm nhạc có bản quyền, từ đó hỗ trợ các nền tảng âm nhạc, nhà phát triển và các tổ chức bảo vệ bản quyền trong việc quản lý và bảo vệ quyền lợi tác giả. Các kết quả từ nghiên cứu có thể giúp nâng cao sự chính xác trong việc phát hiện vi phạm bản quyền, giảm thiểu các tranh chấp liên quan đến quyền sở hữu trí tuệ trong ngành công nghiệp âm nhạc.

GIỚI THIỆU

Hệ thống nhận diện âm nhạc là một phần quan trọng trong việc quản lý và bảo vệ bản quyền âm nhạc trong môi trường số. Trong bối cảnh số lượng bài hát ngày càng tăng và việc phát hiện vi phạm bản quyền trở nên phức tạp hơn, các phương pháp truyền thống cần được tối ưu hóa để đảm bảo tính hiệu quả và tiết kiệm tài nguyên. Dự án nghiên cứu này hướng tới việc ứng dụng các kỹ thuật học máy trong việc nhận diện và dự đoán bản quyền âm nhạc, với mục tiêu cung cấp một giải pháp tự động và hiệu quả cho việc phát hiện vi phạm bản quyền trong ngành công nghiệp âm nhạc [1]. Một trong những thách thức lớn trong việc quản lý bản quyền âm nhạc là việc xác định liệu một bài hát có bị vi phạm bản quyền hay không, khi mà các tác phẩm âm nhạc ngày càng đa dạng và có sự trùng lặp về âm thanh giữa các bài hát. Để giải quyết vấn đề này, dự án sử dụng Chromaprints, một kỹ thuật mã vân tay âm thanh, để trích xuất các đặc trưng đặc thù từ từng bản nhạc [2]. Chromaprints giúp chuyển đổi các đặc điểm âm thanh thành các mã vân tay duy nhất, có thể dễ dàng so sánh và đối chiếu với các bài hát khác trong cơ sở dữ liệu. Các mã vân tay này sẽ được sử dụng như các đặc trưng đầu vào cho mô hình học máy. Thuật toán Random Forest được chọn làm phương pháp phân loại trong nghiên cứu này, nhờ vào khả năng xử lý và phân tích các tập dữ liệu lớn và phức tạp. Random Forest giúp xây dựng một mô hình học sâu mạnh mẽ, có thể phân biệt giữa các bài hát có bản quyền và không có bản quyền, từ đó dự đoán khả năng vi phạm bản quyền của một bài hát mới dựa trên các đặc trưng âm thanh đã được trích xuất. Mục tiêu của dự án không chỉ là phát triển một công cụ hỗ trợ trong việc nhận diện bản quyền âm nhạc mà còn là một bước tiến trong việc ứng dụng các phương pháp học máy vào thực tiễn, góp phần nâng cao hiệu quả trong công tác bảo vệ quyền sở hữu trí tuệ trong ngành âm nhạc. Các kết quả thu được từ nghiên cứu này có thể ứng dụng rộng rãi trong các nền tảng âm nhạc, các dịch vụ trực tuyến, và các tổ chức bảo vệ bản quyền.

Input: Bài hát hoặc link bài hát từ youtube

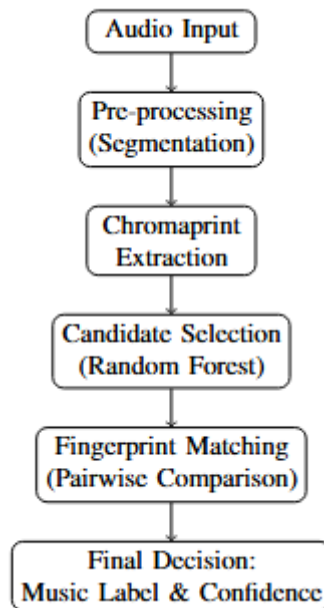
Output: Độ tương đồng của bài hát đó với những bài hát được đăng ký

MỤC TIÊU

Mục tiêu của nghiên cứu này là phát triển một hệ thống tự động nhận diện và phân loại các bài hát có bản quyền, giúp tối ưu hóa quá trình quản lý bản quyền âm nhạc. Để đạt được mục tiêu này, nghiên cứu sử dụng kỹ thuật mã vân tay âm thanh Chromaprints để trích xuất các đặc trưng âm thanh đặc thù từ từng bản nhạc. Các mã vân tay này sẽ được sử dụng làm đầu vào cho mô hình học máy Random Forest, nhằm phân tích và dự đoán khả năng vi phạm bản quyền của các bài hát. Việc tự động hóa quá trình này không chỉ giúp giảm thiểu chi phí và thời gian mà còn nâng cao độ chính xác trong việc phát hiện các trường hợp vi phạm bản quyền. Mục tiêu tiếp theo là tối ưu hóa quá trình trích xuất đặc trưng âm thanh từ Chromaprints. Bằng cách cải thiện độ chính xác và chất lượng của các mã vân tay, hệ thống sẽ có khả năng phân biệt chính xác các bài hát, giảm thiểu khả năng xảy ra sai sót trong việc xác định các bản nhạc trùng lặp hoặc có vi phạm bản quyền. Điều này đóng vai trò quan trọng trong việc nâng cao hiệu suất và độ tin cậy của mô hình học máy, giúp hệ thống có thể xử lý các tập dữ liệu lớn và phức tạp trong thực tế. Cuối cùng, nghiên cứu đặt mục tiêu đánh giá và cải thiện hiệu quả của mô hình Random Forest trong việc phân loại các bài hát có bản quyền và không có bản quyền. Nghiên cứu sẽ tiến hành đánh giá độ chính xác, độ nhạy và độ đặc hiệu của mô hình trong nhiều tình huống khác nhau để tối ưu hóa các tham số mô hình. Mục tiêu này không chỉ chứng minh tính khả thi của mô hình trong việc nhận diện bản quyền âm nhạc mà còn đóng góp vào việc phát triển các công cụ bảo vệ quyền sở hữu trí tuệ trong ngành công nghiệp âm nhạc, góp phần nâng cao hiệu quả công tác quản lý bản quyền và bảo vệ quyền lợi của tác giả.

NỘI DUNG VÀ PHƯƠNG PHÁP

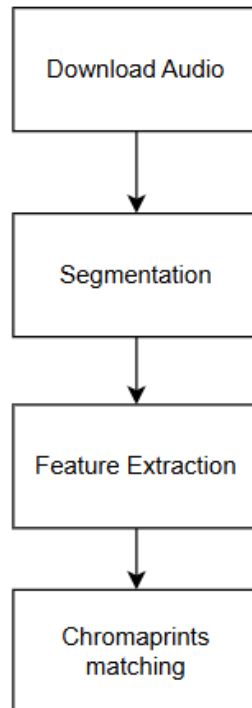
Nghiên cứu này hướng tới việc phát triển một hệ thống tự động nhận diện bản quyền âm nhạc, sử dụng các kỹ thuật tiên tiến trong lĩnh vực mã vân tay âm thanh và học máy. Mục tiêu chính là xây dựng một giải pháp có thể tự động nhận diện và phân loại các bài hát có bản quyền từ nhiều nguồn âm nhạc khác nhau, bao gồm cả các tệp âm thanh cục bộ và video YouTube.



Hình 1: Kiến trúc của hệ thống

Bước 1: Tiền xử lý và trích xuất đặc trưng sơ bộ

Để chuẩn hóa và phân đoạn dữ liệu âm thanh, thư viện Essentia [5] sẽ được sử dụng, cung cấp các công cụ mạnh mẽ để trích xuất các đặc trưng âm thanh quan trọng. Các tệp âm thanh, khoảng 2000 bài hát sẽ được thu thập từ nguồn cục bộ và công cụ Youtube yt-dlp, và sẽ được chuyển đổi thành các đoạn nhỏ (dữ liệu được chuẩn hóa về sampling rate 44.1Khz, khử nhiễu và chia thành các khung 10 giây mỗi đoạn với độ chồng lấn 5 giây), để trích xuất vector đặc trưng 512 chiều, từ đó trừu tượng hóa các yếu tố hài hòa và nhịp điệu quan trọng cho bước phân tích tiếp theo, giúp giữ lại các đặc điểm âm nhạc quan trọng trong suốt quá trình phân tích. Quá trình tiền xử lý này giúp giảm thiểu sự mất mát thông tin trong khi chuẩn bị dữ liệu cho bước tiếp theo.

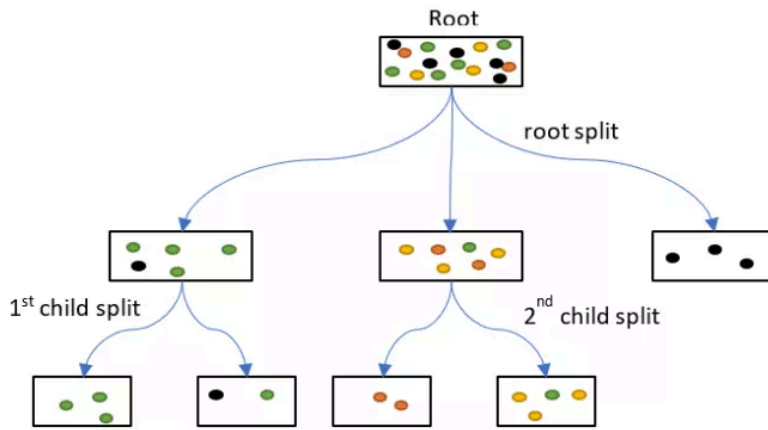


Hình 2: Luồng xử lý Audio

Bước 2: Tạo Chromaprint cho âm thanh

Sau khi tiền xử lý, Chromaprint sẽ được áp dụng để tạo ra mã vân tay âm thanh (audio fingerprints) cho mỗi bài hát. Chromaprint sẽ sinh chuỗi vân tay với ngưỡng threshold = 0,4, giúp chuyển đổi các đặc trưng âm thanh của bài hát thành một chuỗi nhị phân, với mỗi chuỗi đại diện cho các đặc điểm hài hòa của bài hát đó. Điều này giúp nhận diện bài hát một cách chính xác và nhanh chóng. Các mã vân tay này sẽ trở thành đầu vào cho mô hình Random Forest, một thuật toán học máy mạnh mẽ và hiệu quả. Mô hình này được huấn luyện với dữ liệu âm nhạc đã được trích xuất các đặc trưng, giúp phân loại các bài hát và dự đoán khả năng vi phạm bản quyền.

Bước 3: Phân loại bằng Random Forest Classifier



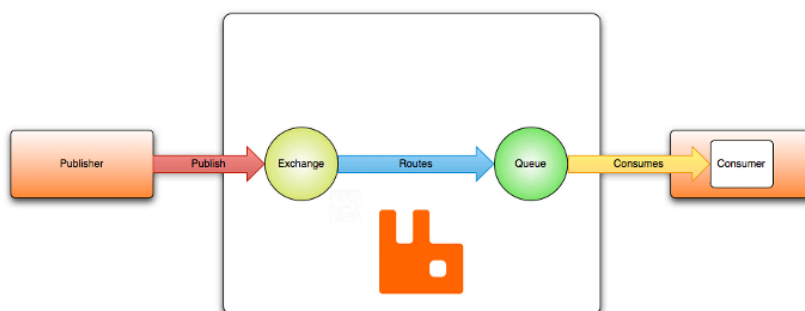
Hình 3: Kiến trúc mô hình Random Forest Classifier

Để mở rộng khả năng nhận diện từ các nguồn âm nhạc trực tuyến, hệ thống sẽ sử dụng công cụ youtube-dl hoặc yt-dlp để tải video từ YouTube, từ đó trích xuất âm thanh và áp dụng Chromaprint để phân tích và nhận diện các bài hát có trong video. Việc tích hợp với các nền tảng video trực tuyến giúp hệ thống có thể nhận diện âm nhạc không chỉ từ các tệp âm thanh cục bộ mà còn từ các video phổ biến trên internet, một tính năng rất quan trọng trong việc bảo vệ bản quyền âm nhạc trong môi trường số hiện nay.

Bước 4: Tối ưu hệ thống với RabbitMQ

Để tối ưu hóa hiệu suất và khả năng mở rộng của hệ thống, RabbitMQ sẽ được tích hợp vào quy trình xử lý. RabbitMQ giúp quản lý hàng đợi và phân phối các công việc bất đồng bộ, đảm bảo hệ thống có thể xử lý đồng thời nhiều tác vụ mà không ảnh hưởng đến hiệu suất. Điều này đặc biệt quan trọng khi hệ thống phải xử lý lượng dữ liệu lớn từ nhiều nguồn khác nhau, đồng thời duy trì khả năng phản hồi nhanh chóng và hiệu quả.

"Hello, world" example routing



Hình 4: Kiến trúc RabbitMQ

Phương pháp này không chỉ đảm bảo tính chính xác trong việc nhận diện bản quyền âm nhạc mà còn giúp hệ thống trở nên linh hoạt và dễ dàng mở rộng trong môi trường thực tế. Hệ thống sẽ có khả năng xử lý dữ liệu âm nhạc quy mô lớn, giảm thiểu chi phí và thời gian so với các phương pháp truyền thống, đồng thời duy trì hiệu suất cao trong các ứng dụng cần xử lý dữ liệu thời gian thực. Với sự kết hợp của Chromaprint, Random Forest, RabbitMQ, và các công cụ hỗ trợ khác, nghiên cứu này sẽ đóng góp vào việc xây dựng các công cụ hiệu quả và tiết kiệm chi phí cho việc quản lý và bảo vệ bản quyền âm nhạc trong ngành công nghiệp âm nhạc số.

KẾT QUẢ MONG ĐỢI

Hệ thống nhận diện bản quyền âm nhạc đề xuất sẽ cung cấp khả năng nhận diện chính xác và hiệu quả các bài hát có bản quyền từ các file âm nhạc cục bộ và video YouTube. Bằng cách sử dụng các kỹ thuật mã vân tay âm thanh tiên tiến, hệ thống sẽ đảm bảo độ chính xác cao trong việc nhận diện và phân loại nội dung âm nhạc. Bên cạnh đó, việc tích hợp RabbitMQ sẽ tối ưu hóa luồng công việc của hệ thống bằng cách quản lý các tác vụ bất đồng bộ, đảm bảo khả năng mở rộng và nâng cao hiệu suất, đặc biệt khi xử lý một lượng lớn dữ liệu trong thời gian thực. Cách tiếp cận này không chỉ tăng tốc quá trình nhận diện mà còn giúp hệ thống có thể xử lý nhiều yêu cầu đồng thời một cách mượt mà, phù hợp với các ứng dụng quy mô lớn.

TÀI LIỆU THAM KHẢO

[1] A. Wang, "An Industrial-Strength Audio Search Algorithm," in Proc. 4th Int. Conf.

Music Information Retrieval (ISMIR), 2003.

[2] J. Haitsma and T. Kalker, "A Highly Robust Audio Fingerprinting System," in Proc. 3rd Int. Conf. Music Information Retrieval (ISMIR), 2002.

[3] P. Cano, E. Batle, T. Kalker, and J. Haitsma, "A Review of Audio Fingerprinting," in Proc. Int. Conf. Music Information Retrieval (ISMIR), 2005.

[4] G. Tzanetakis and P. Cook, "Musical Genre Classification of Audio Signals," IEEE Trans. Speech Audio Process., vol. 10, no. 5, pp. 293–302, Jul. 2002.

[5] A. Bogdanov, B. W. Klapuri, and E. Gómez, "Essentia: An Audio Analysis Library for Music Information Retrieval," in Proc. 21st ACM Int. Conf. Multimedia, 2013, pp. 1127–1130.

[6] M. Casey, R. Veltkamp, M. Goto, M. Leman, C. Rhodes, and M. Slaney, "Content-Based Music Information Retrieval: Current Directions and Future Challenges," Proc. IEEE, vol. 96, no. 4, pp. 668–696, Apr. 2008.

[7] B. S. N. Reddy, B. S. Venkata, I. Manohar, S. Abhishek, and Ariel, "Aural Signatures: Audio Fingerprinting Techniques for Real-Time Audio Recognition," Proc. 2024 3rd Int. Conf. on Sentiment Analysis and Deep Learning (ICSADL), Bhimdatta, Nepal, 25–27 Jul. 2024.

[8] M. A. Casey, R. C. Veltkamp, M. Goto, M. Leman, C. Rhodes, and M. Slaney, "Content-based music information retrieval: current directions and future challenges," Proc. IEEE, vol. 96, no. 4, Apr. 2008.

[9] A. P. Patil, L. J. Itagi, A. C. S., A. G., and M. Ravi, "Design and Implementation of an Audio Fingerprinting System for the Identification of Audio Recordings," in Proc. 2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC), 2021.

[10] N. Borkar, S. Patre, R. S. Khalsa, R. Kawale, and P. Chakkurkar, "Music Plagiarism Detection using Audio Fingerprinting and Segment Matching," in Proc. 2021 Smart Technologies, Communication and Robotics (STCR), 2021.

[11] T. Toshniwal, P. Tandon, and N. P., "Music Genre Recognition Using Short Time Fourier Transform and CNN," in Proc. 2022 International Conference on Computer Communication and Informatics (ICCCI), 2022.

- [12] A. Elbir, H. O. Ilhan, G. Serbes, and N. Aydin, "Short Time Fourier Transform based music genre classification," in Proc. 2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT), 2018.
- [13] S. Prakash and S. Kiran, "Obtain Better Accuracy Using Music Genre Classification System on GTZAN Dataset," in Proc. 2022 IEEE North Karnataka Subsection Flagship International Conference (NKCon), 2022.
- [14] B. Chapman-Krish, "Exploring Deep Learning Techniques using SciKit-Learn and TensorFlow for Music Genre Classification on the FMA Dataset," 2022.
- [15] B. Chen, "Music Audio Rhythm Recognition Based on Recurrent Neural Network," Mobile Information Systems, vol. 2022, Article ID 6249798, 2022.
- [16] M. Chikanbanjar, "Comparative analysis between audio fingerprinting algorithms," Int. J. Comput. Sci. Eng. Technol. May 2017.