ĐẠI HỌC ĐÀ NẴNG

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

**VÀ TRUYỀN THÔNG VIỆT - HÀN**

****

**BÁO CÁO TỔNG KẾT  
ĐỀ TÀI NGHIÊN CỨU KHOA HỌC CỦA SINH VIÊN   
NĂM 2022-2023**

**GENERATE MUSIC IN TIME SERIES WITH LSTM**

**ĐHVH-2023-SV-09**

Thuộc lĩnh vực khoa học và công nghệ: Công nghệ thông tin

(chuyên ngành Trí tuệ nhân tạo và Khoa học dữ liệu)

Sinh viên chịu trách nhiệm thực hiện đề tài: Phan Lê Việt Hùng, Phạm Văn Mỹ

Ngành học: Công nghệ thông tin

(chuyên ngành Trí tuệ nhân tạo và Khoa học dữ liệu)

Người hướng dẫn: TS. Đặng Đại Thọ

**Đà Nẵng, tháng 07 năm 2023**

ĐẠI HỌC ĐÀ NẴNG

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

**VÀ TRUYỀN THÔNG VIỆT – HÀN**

**BÁO CÁO TỔNG KẾT**

**ĐỀ TÀI NGHIÊN CỨU KHOA HỌC CỦA SINH VIÊN**

**NĂM 2022-2023**

**GENERATE MUSIC IN TIME SERIES WITH LSTM**

**ĐHVH-2023-SV-09**

| **Xác nhận của đơn vị chủ trì**  *(ký, họ tên, đóng dấu)* | **Sinh viên chịu trách nhiệm chính**  **(***ký, họ tên)* |
| --- | --- |

**Đà Nẵng, tháng 07 năm 2023**

**MỤC LỤC**

[**LỜI CẢM ƠN 4**](#_heading=h.wd7rycp3znje)

[**THÔNG TIN KẾT QUẢ NGHIÊN CỨU CỦA ĐỀ TÀI 6**](#_heading=h.3znysh7)

[**CHƯƠNG 1: TỔNG QUAN 12**](#_heading=h.3pgnjuxcq7su)

[1.1. Tổng quan tình hình nghiên cứu ở trong nước và nước ngoài 12](#_heading=h.ej3x9so63kn2)

[1.2. Lý do chọn đề tài 12](#_heading=h.2gio97paykqm)

[1.3. Mục tiêu đề tài 12](#_heading=h.hvhrcrvdnva3)

[1.4. Phương pháp nghiên cứu 13](#_heading=h.v8yk8wibif5)

[1.5. Đối tượng và phạm vi nghiên cứu. 13](#_heading=h.iwvxmgiui6n2)

[**CHƯƠNG 2: XỬ LÝ NGÔN NGỮ TỰ NHIÊN 14**](#_heading=h.4t7g2z40ch8d)

[2.1. Tổng quan về kỹ thuật Xử lý ngôn ngữ tự nhiên 14](#_heading=h.mk5wxsk3su12)

[2.1.1. Khái niệm về Xử lý ngôn ngữ tự nhiên 14](#_heading=h.og08pq4uupm6)

[2.1.2. Quy trình Xử lý ngôn ngữ tự nhiên 14](#_heading=h.hawfw3ygz39a)

[2.2. Các thành phần của Xử lý ngôn ngữ tự nhiên 15](#_heading=h.9zq6i7ombc68)

[2.3. Các phương pháp Xử lý ngôn ngữ tự nhiên 15](#_heading=h.926w31bdy820)

[2.4. Ứng dụng Xử lý ngôn ngữ tự nhiên 16](#_heading=h.hx05ypundilu)

[2.5. Các thách thức trong xử lý ngôn ngữ tự nhiên 17](#_heading=h.66a05wqlgdlv)

[2.6. Cơ sở dữ liệu Từ tục tĩu 18](#_heading=h.p0j3iwwalg51)

[2.6.1. Sơ lược về Từ tục tĩu 18](#_heading=h.j66pmydmthlf)

[2.6.2. Tập dữ liệu từ tục tĩu 18](#_heading=h.5v2lcegiqoul)

[**CHƯƠNG 3: ĐỀ XUẤT GIẢI PHÁP 19**](#_heading=h.gjx88z6ptbst)

[3.1. Data Preprocessing 19](#_heading=h.w14jeey29qvf)

[3.2. Word Embedding 20](#_heading=h.vdxd5hzix0b)

[3.2.1. Word2Vec 20](#_heading=h.48vd1yqeohmg)

[3.2.2. BERT, RoBERTa và PhoBERT 21](#_heading=h.w4ipqanposk4)

[3.2.2.1. BERT 21](#_heading=h.vor59gxe6d4z)

[3.2.2.2 RoBERTa và PhoBERT: 24](#_heading=h.iv7pi8mj9fag)

[3.3. Machine Learning và Deep Learning 24](#_heading=h.nci0fg1h7j3a)

[3.3.1. Thuật toán SVM 24](#_heading=h.ady5zu94n2m6)

[3.3.2. Thuật toán CNN 25](#_heading=h.tq6dz5j3d884)

[3.3.3. Mạng LSTM và BiLSTM 26](#_heading=h.7hruco17qxtz)

[**CHƯƠNG 4: THỰC NGHIỆM VÀ ĐÁNH GIÁ 28**](#_heading=h.xvi7o3evwoe5)

[4.1. Xây dựng bộ dữ liệu 28](#_heading=h.2838tg45juun)

[4.2. Tiền xử lý dữ liệu 28](#_heading=h.qdjenxrmn31h)

[4.3. Embed dữ liệu 29](#_heading=h.sly2ngvuirca)

[4.4. Xây dựng mô hình học máy/học sâu 29](#_heading=h.r6n6ihx93vb4)

[4.5. Kết quả nghiên cứu 31](#_heading=h.fwuzxten7hnh)

[4.6. Kết luận 38](#_heading=h.8rlkxx26j6jc)

[**KẾT LUẬN VÀ KIẾN NGHỊ 39**](#_heading=h.l4973rooxvpv)

[5.1 Kết luận 39](#_heading=h.ms0m90zak7k2)

[5.2 Kiến nghị 40](#_heading=h.5oyjkk7ocsxe)

[**TÀI LIỆU THAM KHẢO 40**](#_heading=h.yp4duip3hdj4)

**DANH MỤC HÌNH ẢNH**

Hình 3.1 Minh hoạ phương pháp CBOW 16

Hình 3.2 Minh hoạ thuật toán SVM 18

Hình 3.3 Minh hoạ thuật toán CNN 19

Hình 3.4 Minh hoạ thuật toán RoBERTa 21

Hình 3.5 Minh hoạ thuật toán PhoBERT 22

**DANH MỤC BẢNG BIỂU**

Bảng 4.1 Kết quả so sánh giữa Word2Vec và PhoBERT 23

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# LỜI CẢM ƠN

Chúng em muốn dành cho thầy Nguyễn Hữu Nhật Minh những lời cảm ơn chân thành vì kiến thức và kinh nghiệm chuyên môn của thầy đã giúp chúng em nâng cao hiểu biết về chủ đề của bài báo cáo. Thầy đã dành thời gian để hướng dẫn và chia sẻ những góp ý, định hướng giúp chúng tôi hiểu rõ hơn về cách thực hiện nghiên cứu và phân tích dữ liệu.

Thêm vào đó, chúng em cũng muốn bày tỏ lòng biết ơn đối với sự tận tâm và sự nhiệt tình của thầy trong việc hỗ trợ và giải đáp các câu hỏi của chúng em. Thầy luôn sẵn sàng dành thời gian để lắng nghe và đưa ra những lời khuyên hữu ích, giúp chúng em hoàn thiện bài báo cáo của mình.

Chúng em xin chân thành cảm ơn!

Nhóm sinh viên thực hiện

Lê Cẩm Bằng

Phan Lê Việt Hùng

| ĐẠI HỌC ĐÀ NẴNG  **TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG VIỆT - HÀN** | **CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM**  **Độc lập – Tự do – Hạnh phúc**    *Đà Nẵng, ngày 25 tháng 03 năm 2023* |
| --- | --- |

# THÔNG TIN KẾT QUẢ NGHIÊN CỨU CỦA ĐỀ TÀI

**1. Thông tin chung:**

- Tên đề tài: Nghiên cứu sử dụng trí tuệ nhân tạo để phát hiện website có chứa nội dung thô tục

- Mã đề tài: ĐHVH-2023-SV-09

- Sinh viên chịu trách nhiệm: Lê Cẩm Bằng, Phan Lê Việt Hùng

- Lớp: 21AD Khoa: Khoa học máy tính Năm thứ: 2 Số năm đào tạo: 4,5 năm

- Người hướng dẫn: TS. Đặng Đại Thọ

- Thời gian thực hiện: tháng 10 năm 2022

**2. Mục tiêu :**

Đề tài tập trung nghiên cứu, phát triển hệ thống phát hiện các website/ mạng xã hội có chứa nội dung thô tục.

**3. Tính mới và sáng tạo:**

Với lượng dữ liệu khổng lồ có sẵn trên internet, việc xem xét thủ công từng trang web để lọc nội dung không phù hợp là một nhiệm vụ khó khăn. Việc sử dụng các thuật toán AI có thể xác định, phân loại hiệu quả và chính xác các trang web chứa nội dung thô tục, giúp ngăn chặn người dùng truy cập vào các trang web có nội dung thô tục một cách dễ dàng hơn. Trong tương lai, các thuật toán mới được ra mắt và cải thiện sẽ nâng cao hơn nữa độ chính xác và hiệu quả trong việc phát hiện nội dung thô tục.

**4. Kết quả nghiên cứu:**

* Một phương pháp dựa vào AI để phát hiện những trang thông tin có nội dung thô tục.
* Một báo cáo tổng kết những lý thuyết và thực nghiệm trong quá trình nghiên cứu
* Một bài báo công bố trên một tạp chí khoa học quốc tế về việc ứng dụng công nghệ AI vào lĩnh vực giám sát nội dung thô tục trên internet.

**5. Sản phẩm:**

* 01 sản phẩm ứng dụng
* 01 báo cáo
* 01 bài báo khoa học quốc tế

**6. Về các đóng góp của đề tài cho giáo dục và đào tạo, kinh tế - xã hội, an ninh, quốc phòng:**

* Báo cáo, bài báo của đề tài góp phần phục vụ giáo dục, đào tạo.
* Phương pháp và sản phẩm phần mềm của đề tài góp phần vào việc ngăn chặn các website có nội dung thô tục

**7. Công bố khoa học của sinh viên từ kết quả nghiên cứu của đề tài** (ghi rõ họ tên tác giả, nhan đề và các yếu tố về xuất bản nếu có) hoặc nhận xét, đánh giá của cơ sở đã áp dụng các kết quả nghiên cứu (nếu có):

* Họ tên tác giả: Dai Tho Dang, Cong Phap Huynh, Cam Bang Le, Luong Vuong Nguyen, Si Thin Nguyen (Đặng Đại Thọ, Huỳnh Công Pháp, Lê Cẩm Bằng, Nguyễn Lương Vương, Nguyễn Sĩ Thìn)
* Tên bài báo: Machine Learning Techniques for Vietnamese Profanity Detection.
* Hội nghị: International Conference on Human-centered Artificial Intelligence 2022.

<>

**Sinh viên chịu trách nhiệm chính**

**thực hiện đề tài**

(ký, họ và tên)

**Nhận xét của người hướng dẫn về những đóng góp khoa học của sinh viên thực hiện đề tài** (phần này do người hướng dẫn ghi):

**Xác nhận của đơn vị Người hướng dẫn**

*(Ký, ghi rõ họ tên)*  *(Ký, ghi rõ họ tên)*

| ĐẠI HỌC ĐÀ NẴNG  **TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG VIỆT - HÀN** | **CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM**  **Độc lập – Tự do – Hạnh phúc**    *Đà Nẵng, ngày 25 tháng 03 năm 2023* |
| --- | --- |

**THÔNG TIN VỀ SINH VIÊN  
CHỊU TRÁCH NHIỆM CHÍNH THỰC HIỆN ĐỀ TÀI**

**I. SƠ LƯỢC VỀ SINH VIÊN:**

Họ và tên: Lê Cẩm Bằng

Sinh ngày: 08 tháng 11 năm 2003

Nơi sinh: Đà Nẵng

Lớp: 21AD Khóa: 2021-2026

Khoa: Khoa học máy tính

Địa chỉ liên hệ: đường Phạm Như Xương - Điện Ngọc - Điện Bàn - Quảng Nam

Điện thoại: 0705991986 Email: banglc.21ad@vku.udn.vn

**II. QUÁ TRÌNH HỌC TẬP** (kê khai thành tích của sinh viên từ năm thứ 1 đến năm đang học):

*\* Năm thứ 1:*

Ngành học: Công nghệ thông tin

(chuyên ngành Trí Tuệ Nhân Tạo và Khoa Học Dữ Liệu)

Khoa: Khoa Học Máy Tính

Kết quả xếp loại học tập: Giỏi

Sơ lược thành tích: Giải khuyến khích cuộc thi Best Web Design 2022, Giải khuyến khích cuộc thi Ươm mầm ý tưởng khởi nghiệp và sáng tạo PISI 2022

*\* Năm thứ 2:*

Ngành học: Khoa:

Kết quả xếp loại học tập:

Sơ lược thành tích:

**Xác nhận của đơn vị** **Sinh viên chịu trách nhiệm chính**

**thực hiện đề tài**

*(ký, họ và tên)*

| ĐẠI HỌC ĐÀ NẴNG  **TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG VIỆT - HÀN** | **CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM**  **Độc lập – Tự do – Hạnh phúc**    *Đà Nẵng, ngày 25 tháng 03 năm 2023* |
| --- | --- |

**THÔNG TIN VỀ SINH VIÊN  
CHỊU TRÁCH NHIỆM CHÍNH THỰC HIỆN ĐỀ TÀI**

**I. SƠ LƯỢC VỀ SINH VIÊN:**

Họ và tên: Phan Lê Việt Hùng

Sinh ngày: 18 tháng 08 năm 2003

Nơi sinh: Đà Nẵng

Lớp: 21AD Khóa: 2021-2026

Khoa: Khoa Học Máy Tính

Địa chỉ liên hệ: 126/3 Lê Độ, phường Chính Gián, quận Thanh Khê, T.P. Đà Nẵng

Điện thoại: 0911040205 Email: hungplv.21ad@vku.udn.vn

**II. QUÁ TRÌNH HỌC TẬP** (kê khai thành tích của sinh viên từ năm thứ 1 đến năm đang học):

*\* Năm thứ 1:*

Ngành học: Công nghệ thông tin

(chuyên ngành Trí Tuệ Nhân Tạo và Khoa Học Dữ Liệu)

Khoa: Khoa Học Máy Tính

Kết quả xếp loại học tập: Xuất sắc

Sơ lược thành tích:

Giải khuyến khích cuộc thi Best Web Design 2022,

giải khuyến khích cuộc thi Ươm mầm ý tưởng khởi nghiệp và sáng tạo PISI 2022

*\* Năm thứ 2:*

Ngành học: Khoa:

Kết quả xếp loại học tập:

Sơ lược thành tích:

**Xác nhận của đơn vị** **Sinh viên chịu trách nhiệm chính**

**thực hiện đề tài**

*(ký, họ và tên)*

## 

# 

# CHƯƠNG 1: TỔNG QUAN

## 1.1. Tổng quan tình hình nghiên cứu ở trong nước và nước ngoài

Tại Việt Nam, việc ứng dụng Trí tuệ nhân tạo (AI) trong âm nhạc đang được quan tâm và phát triển. Nhiều nhà nghiên cứu và tổ chức âm nhạc đang khám phá việc kết hợp AI để hỗ trợ các nghệ sĩ tạo ra các bài hát mới và nâng cao trải nghiệm nghe nhạc tổng thể cho người dùng.

Thêm vào đó, việc tích hợp AI vào sáng tác âm nhạc sẽ cho ra đời những ca khúc mới mang đậm phong cách và chất riêng của Việt Nam. Những sáng tác này sẽ góp phần chia sẻ những di sản văn hóa phong phú của Việt Nam với bạn bè cũng như những người đam mê âm nhạc trên toàn thế giới.

Tại nước ngoài, tình hình nghiên cứu ứng dụng AI để sáng tác âm nhạc cũng đang phát triển mạnh mẽ. Các sản phẩm và dự án như Magenta của Google và MuseNet của OpenAI đã góp phần nâng cao khả năng sáng tạo của AI, giúp các nhà sản xuất âm nhạc, nhạc sĩ và nghệ sĩ tối ưu hóa quy trình sáng tạo, rút ngắn thời gian cũng như khối lượng công việc trong sáng tác.

## 1.2. Lý do chọn đề tài

Nhờ sự phát triển như vũ bão của công nghệ hiện nay, AI đã có khả năng phân tích và học hỏi từ lượng lớn dữ liệu âm nhạc từ khắp nơi trên thế giới. Nhờ đó, nó có thể hỗ trợ chúng ta trong quá trình sáng tác, giúp các nghệ sĩ tiết kiệm thời gian và năng lượng để tập trung vào khía cạnh sáng tạo của một tác phẩm. Theo xu hướng này, chúng em đã lựa chọn đề tài này nhằm đóng góp vào việc xây dựng trải nghiệm âm nhạc mới cho người dùng, trải nghiệm âm nhạc cá nhân hóa. Chúng em hướng đến việc tạo ra một sản phẩm giúp người dùng có thể tận dụng AI để sáng tác và tận hưởng những bài hát mang phong cách riêng của họ.

## 1.3. Mục tiêu đề tài

Mục tiêu của đề tài là xây dựng một mô hình sáng tạo âm nhạc bằng cách sử dụng LSTM (Long Short-Term Memory). Mô hình này sẽ có khả năng học và hiểu các đặc điểm âm nhạc từ bài hát đầu vào của người dùng và tự động tạo ra bài hát đầu ra mới và đa dạng. Trong tương lai gần, mô hình sẽ giúp người nghe nhạc được trải nghiệm những bài hát được cá nhân hóa, từ đó nâng cao trải nghiệm âm nhạc của họ.

## 1.4. Phương pháp nghiên cứu

Phương pháp nghiên cứu sử dụng trong đề tài này là kết hợp giữa phương pháp nghiên cứu lý thuyết và phương pháp nghiên cứu thực nghiệm.

Trong phương pháp nghiên cứu lý thuyết, chúng em sẽ tập trung vào các kỹ thuật tiền xử lý dữ liệu dành riêng cho âm nhạc như transpose, mã hóa dữ liệu dưới định dạng chuỗi thời gian, ánh xạ dữ liệu,… và mô hình học sâu LSTM.

Phần nghiên cứu thực nghiệm sẽ tiến hành trên các phương pháp, kỹ thuật đã được nghiên cứu ở phần lý thuyết. Bằng cách áp dụng nguyên tắc của chuỗi thời gian, mô hình sẽ có thể dự đoán các nốt nhạc có thể xuất hiện trong tương lai dựa vào những nốt nhạc hiện tại.

## 1.5. Đối tượng và phạm vi nghiên cứu.

Đối tượng nghiên cứu của đề tài là cách các mô hình trí tuệ nhân tạo hiểu về nhạc lý và ứng dụng chúng trong việc sáng tác âm nhạc.

Phạm vi nghiên cứu sẽ bao gồm các phương pháp và công nghệ để phân tích, xử lý và sáng tạo âm nhạc nói riêng và âm thanh nói chung.

# CHƯƠNG 2: NHẠC LÝ ỨNG DỤNG TRONG NGHIÊN CỨU

Về cốt lõi, âm nhạc được hình thành bằng cách kết hợp các nốt nhạc và các khoảng dừng trong một tác phẩm. Các yếu tố cơ bản của nốt nhạc là cao độ và trường độ nốt nhạc (hay độ dài của nốt nhạc).

## 2.1. Cao độ, trường độ và số chỉ nhịp

Cao độ trong âm nhạc là một khái niệm để miêu tả các âm thanh hoặc nốt nhạc có tần số cao hay thấp. Cao độ đề cập đến tên của ghi chú cùng với quãng tám của nó. Cao độ cao được liên kết với âm thanh hay nốt nhạc có tần số cao, trong khi cao độ thấp liên quan đến âm thanh hay nốt nhạc có tần số thấp. Điều này được đại diện bởi các nốt nhạc trên bảng nhạc, từ các nốt nhạc thấp như C, D, E, F, G, A, B cho đến các nốt nhạc cao hơn như C, D, E, F, G, A, B, và tiếp tục lặp lại. Đây là một trong những yếu tố quan trọng để xác định âm sắc và cảm xúc của bản nhạc.

Trường độ tương ứng với độ dài của một nốt nhạc, được đo bằng phách. Trong ký hiệu âm nhạc tiêu chuẩn, thời lượng được thể hiện bằng các ký hiệu như nốt nguyên, nốt nửa, nốt đen, nốt thứ tám và nốt thứ mười sáu. Ví dụ: nốt thứ mười sáu có thời lượng 0,25 phách.

Chỉ số nhịp là một khía cạnh quan trọng khác của âm nhạc, xác định số lượng nhịp trong mỗi chu kỳ nhạc và loại nhịp mà mỗi nhịp đại diện cho. Từ đó số chỉ nhịp là dấu hiệu giúp nhận biết bản nhạc đó có nhịp 2, nhịp 3 hay nhịp 4. Chỉ số nhịp thường được biểu diễn dưới dạng một ký hiệu ở đầu bản nhạc và thường là một cặp số, thường là hai số được đặt trên đứng nhau. Hai số này thể hiện hai thông tin chính:

Số trên: Thể hiện số lượng nhịp (beat) trong mỗi chu kỳ nhạc (measure). Số trên cho biết số lượng nhịp trong một chu kỳ nhạc mà người chơi hoặc người hát phải đếm.

Số dưới: Thể hiện loại nhịp (beat) mà mỗi nhịp đại diện cho. Số dưới cho biết một nhịp có giá trị bao nhiêu trong một khoảng thời gian nhất định, thường là 1/4, 1/8, 1/16, v.v.

## 2.2. Key, tonic và mode

Key (điệu tính) là một nhóm cao độ tạo thành âm sắc trung tâm của một bản nhạc. Nó được xác định bởi sự kết hợp của chủ âm và điệu. Có tổng cộng 24 phím, được hình thành bởi 12 âm bổ và 2 chế độ. Chẳng hạn, C trưởng và D thứ là những ví dụ về các khóa khác nhau.

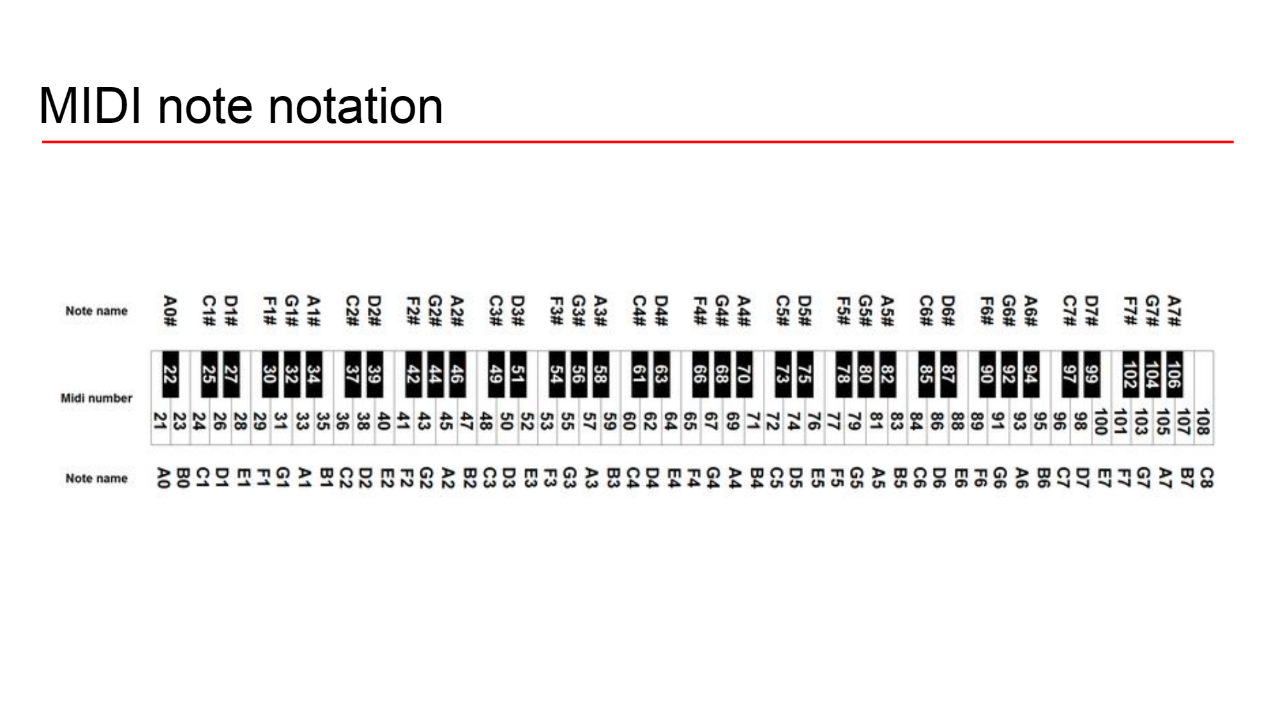
Tonic (chủ âm) đề cập đến nốt chính của một phím, thường đóng vai trò là nền tảng chính hoặc điểm nghỉ ngơi trong một tác phẩm âm nhạc.

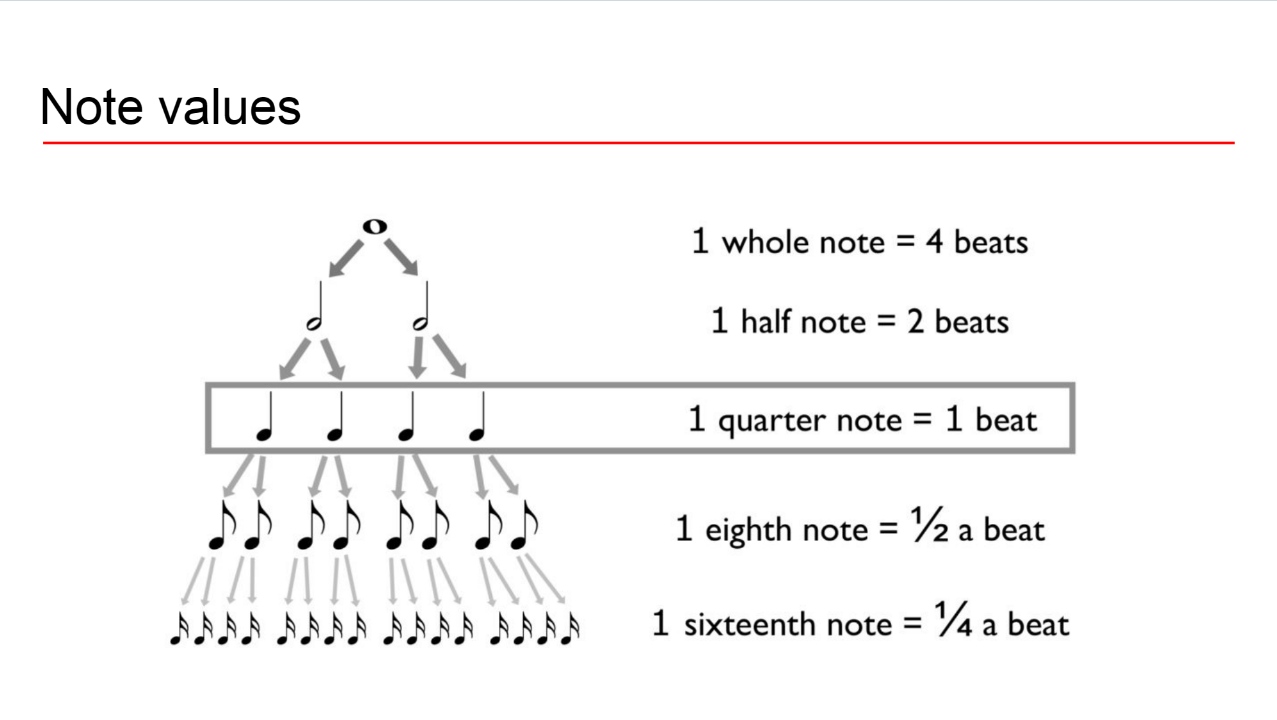
Mode (Điệu thức) bao gồm các thang âm khác nhau, chẳng hạn như âm giai trưởng và âm giai thứ, mang đến những cảm xúc và tâm trạng riêng biệt cho âm nhạc.

Hiểu được các yếu tố nền tảng này là điều cần thiết cho bất kỳ quá trình phân tích và sáng tác âm nhạc nào.

Giai điệu âm nhạc được tạo thành từ những nốt nhạc và dấu nghỉ. Nốt nhạc bao gồm 2 thành phần là pitch và duration.

Pitch là … Note name + octave. Bao gồm C, D, E,...



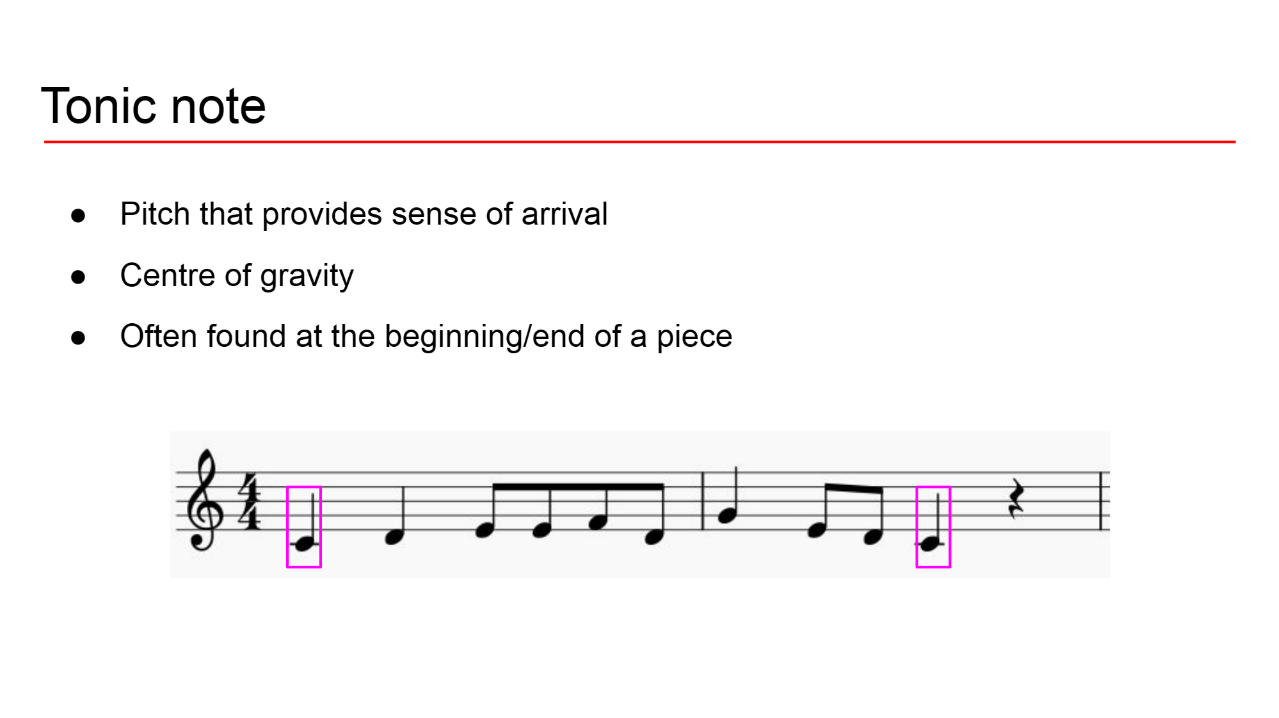
Duration là … Bao gồm 0.25 beats (sixteenth note),...

Time signature là … Bao gồm nhịp 4/4, 2/4,...

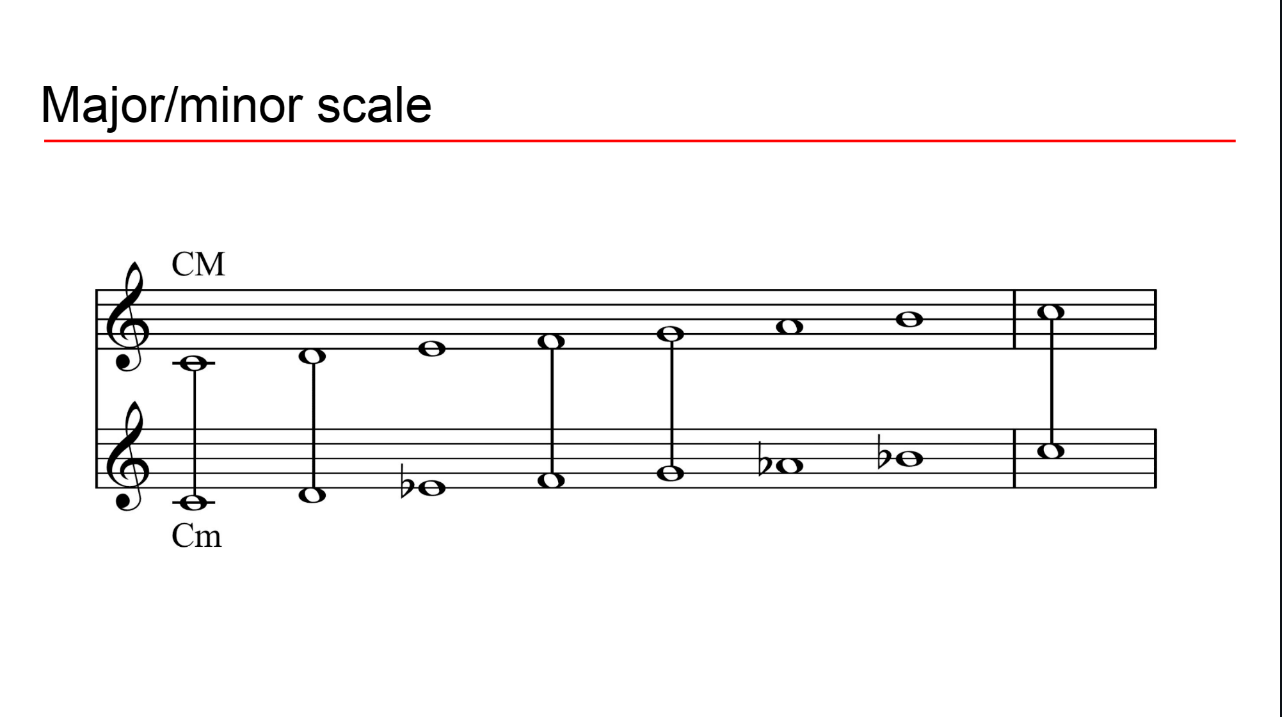
Key is Group of pitches (i.e., sca/e) that forms the centre of a piece.

Key = Tonic + mode. There are 24 keys (12 tonics \* 2 modes). E.g.,C major, D minor,...

Tonic là …



Mode là… Gồm major và minor



# CHƯƠNG 3: DỮ LIỆU VÀ CÔNG CỤ

## 3.1. Định dạng dữ liệu

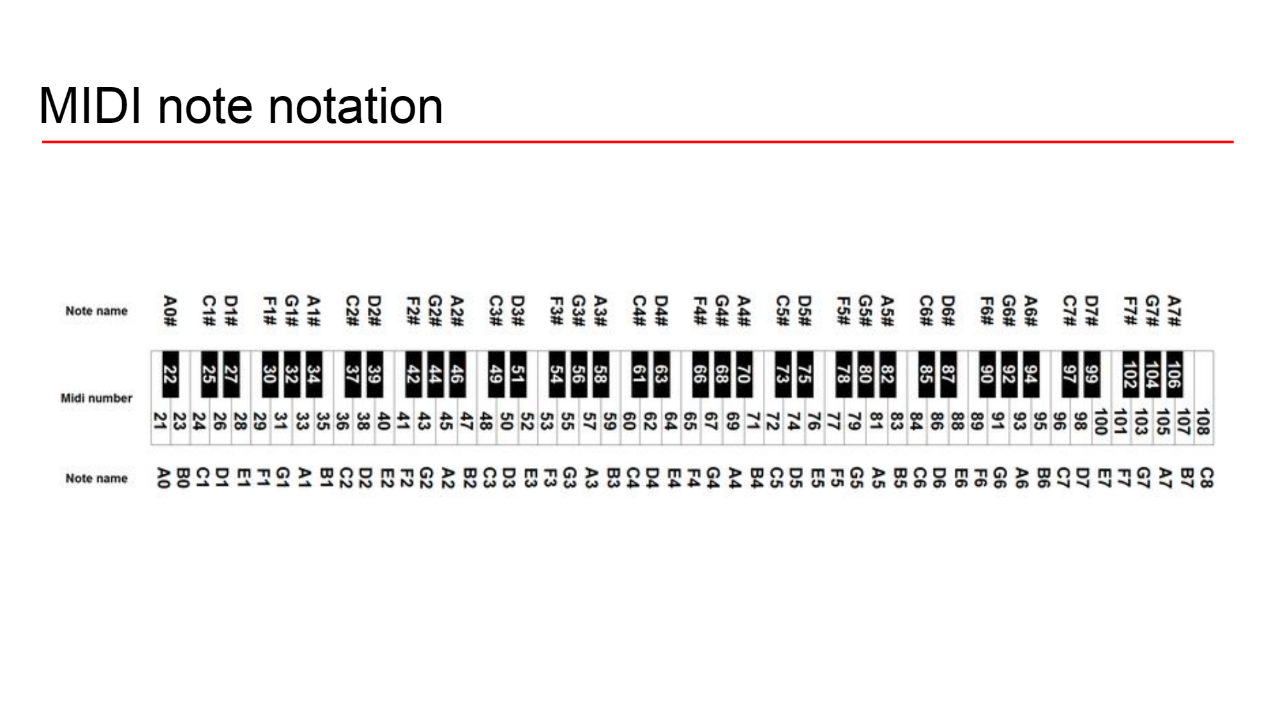
Trong nỗ lực sáng tác và phân tích âm nhạc của mình, chúng tôi sử dụng định dạng dữ liệu kern. Định dạng kern có những ưu điểm và hạn chế nhất định, làm cho nó phù hợp với các tác vụ cụ thể. Một số điểm chính về định dạng dữ liệu kern bao gồm:

Thuận lợi:

* Khả năng đọc: Định dạng kern có thể đọc được bằng con người, cho phép người dùng hiểu dữ liệu âm nhạc một cách dễ dàng.
* Tính linh hoạt: Nó có thể đại diện cho các thuộc tính âm nhạc khác nhau và rất phù hợp cho nghiên cứu và phân tích âm nhạc học.
* Tính linh hoạt: Các tệp Kern có thể được chuyển đổi sang các định dạng khác, cho phép khả năng tương tác với các công cụ phần mềm khác nhau.

Hạn chế:

* Độ phức tạp: Định dạng kern có thể trở nên khó xử lý khi xử lý các tác phẩm có độ phức tạp cao.
* Thiếu tiêu chuẩn hóa: Không giống như các định dạng được sử dụng rộng rãi như MIDI, định dạng kern không được tiêu chuẩn hóa trong ngành công nghiệp âm nhạc.

Để thu thập tập dữ liệu đa dạng và phong phú, chúng tôi chủ yếu lấy dữ liệu từ tập dữ liệu ESAC, bao gồm hơn 5.000 bài hát thuộc nhiều thể loại. Bộ dữ liệu này đảm bảo rằng phân tích của chúng tôi bao gồm nhiều phong cách âm nhạc và giúp chúng tôi đưa ra kết luận chắc chắn. Đầu ra sẽ là file MIDI. 

## 3.2. Công cụ sử dụng

Để theo đuổi phân tích và sáng tác âm nhạc toàn diện, chúng tôi tận dụng các công cụ mạnh mẽ sau:

Thư viện Keras

Keras là một khung học sâu mã nguồn mở đóng vai trò quan trọng trong quá trình phân tích âm nhạc của chúng tôi. Mặc dù ban đầu được thiết kế cho các tác vụ học máy nói chung, Keras đã tìm thấy các ứng dụng trong lĩnh vực âm nhạc với kết quả ấn tượng. Giao diện thân thiện với người dùng và khả năng tích hợp liền mạch với các thư viện khác khiến nó trở thành tài sản không thể thiếu cho nghiên cứu của chúng tôi.

Thư viện Music21

Thư viện Music21 là một lợi ích cho các nhà âm nhạc cũng như các nhà soạn nhạc. Thư viện dựa trên Python này cho phép chúng tôi thao tác và phân tích dữ liệu âm nhạc một cách dễ dàng. Với Music21, chúng tôi có thể thực hiện các tác vụ như phân tích nốt nhạc, phân tích khóa và hợp âm, v.v. Đó là tính linh hoạt và tài liệu phong phú hợp lý hóa quy trình phân tích của chúng tôi và làm phong phú thêm sự hiểu biết của chúng tôi về các tác phẩm âm nhạc.

Phần mềm MuseScore

MuseScore là một phần mềm soạn nhạc mạnh mẽ hỗ trợ việc tạo các bản nhạc. Nó cho phép chúng tôi phiên âm các ý tưởng âm nhạc, thử nghiệm các cách sắp xếp khác nhau và tinh chỉnh các tác phẩm. Giao diện trực quan và bộ tính năng toàn diện của MuseScore cho phép chúng tôi trực quan hóa dữ liệu âm nhạc của mình một cách hiệu quả.

Bằng cách khai thác khả năng của những công cụ đáng chú ý này, chúng tôi có được những hiểu biết có giá trị về sự phức tạp của âm nhạc và nâng cao quá trình sáng tác của chúng tôi.

Tại giao điểm của âm nhạc và công nghệ, nỗ lực nghiên cứu của chúng tôi tập trung vào việc trích xuất các mẫu, hiểu cấu trúc âm nhạc và khám phá các kỹ thuật sáng tác sáng tạo. Việc sử dụng định dạng dữ liệu kern, cùng với sự hỗ trợ của Keras, Music21 và MuseScore, cho phép chúng tôi thực hiện những khám phá đột phá trong thế giới âm nhạc.

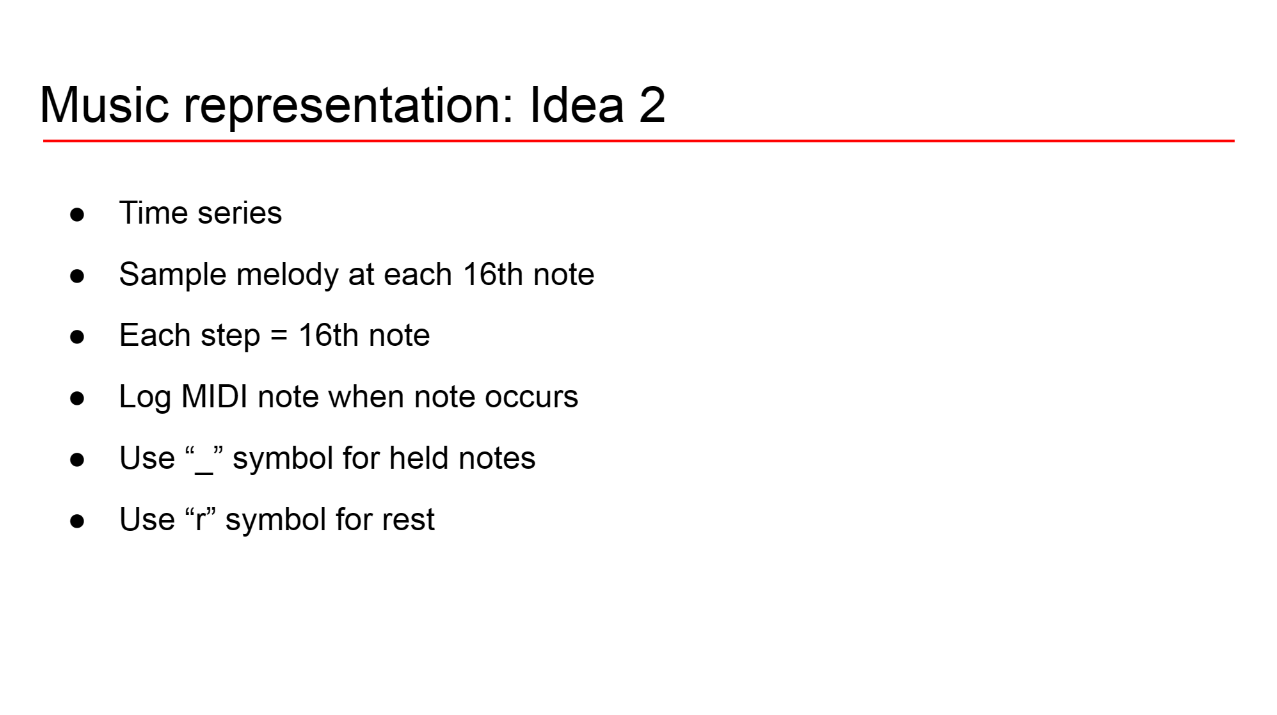
# CHƯƠNG 4: TIỀN XỬ LÝ DỮ LIỆU

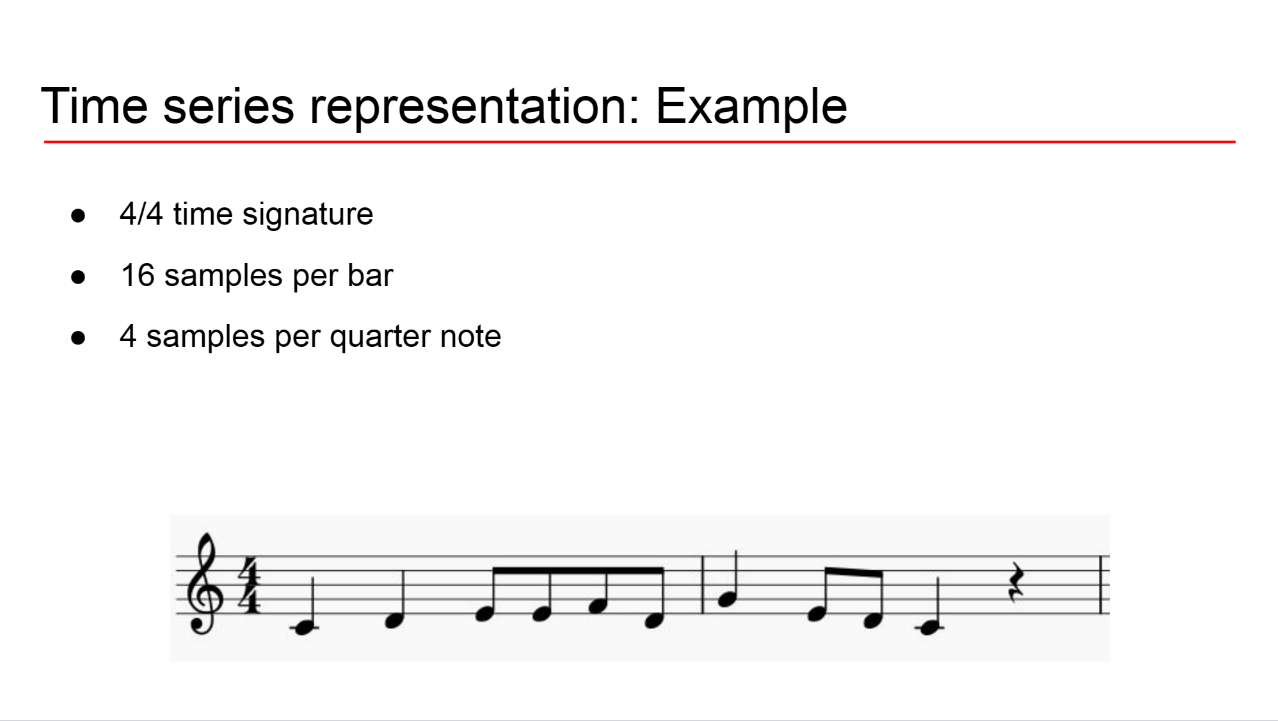
## 4.1. Transpose

* Score: orange
* Part (for different musical instrument): Green
* Measure (bar): Purple
* Note: Blue



## 4.2. Encode data to time series representation







## 4.3. Synthesize dataset into a single data file

## 4.4. Map data to integers with json dictionary

## 4.5. Generate input and output sequences for training

### 

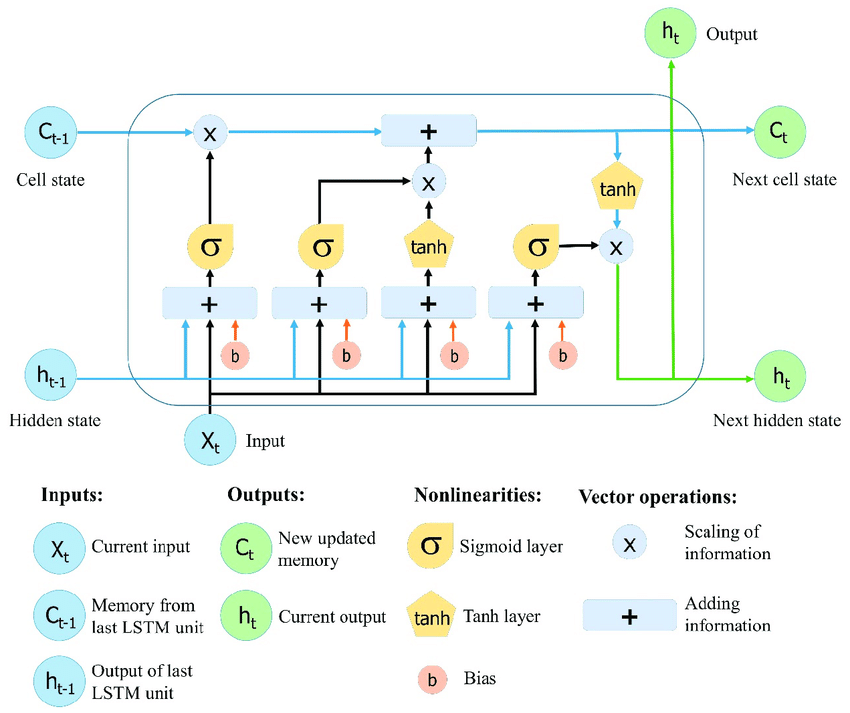
*Hình 2.1 Minh hoạ Quy trình Xử lý ngôn ngữ tự nhiên2*

# 

# CHƯƠNG 5: XÂY DỰNG VÀ HUẤN LUYỆN MÔ HÌNH

## 5.1. Mạng LSTM và BiLSTM

LSTM (Long Short-Term Memory) là một phiên bản mở rộng của mạng RNN (Recurrent Neural Network), được đề xuất vào năm 1997 bởi Sepp Hochreiter và Jürgen Schmidhuber. Nó giải quyết vấn đề phụ thuộc dài hạn của RNN trong đó RNN không thể dự đoán từ được lưu trữ trong bộ nhớ dài hạn nhưng có thể đưa ra dự đoán chính xác hơn từ thông tin gần đây. Trong khi RNN không cung cấp hiệu suất hiệu quả khi khoảng cách tăng lên, LSTM có thể mặc định giữ thông tin trong một thời gian dài. Nó được sử dụng để xử lý, dự đoán và phân loại dựa trên dữ liệu chuỗi thời gian.



Hình Minh hoa thuật toán LSTM

Cite: Le, Xuan Hien & Ho, Hung & Lee, Giha & Jung, Sungho. (2019). Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting. Water. 11. 1387. 10.3390/w11071387.

Một đơn vị LSTM thông thường bao gồm một tế bào (cell), một cổng vào (input gate), một cổng ra (output gate) và một cổng quên (forget gate). Tế bào ghi nhớ các giá trị trong các khoảng thời gian bất ý và ba cổng sẽ điều chỉnh luồng thông tin ra/vào tế bào.

BiLSTM là viết tắt của Bidirectional Long Short-Term Memory, là một mô hình xử lý chuỗi bao gồm hai LSTM: một lấy đầu vào theo hướng tiến và một theo hướng ngược lại². BiLSTMs tăng lượng thông tin có sẵn cho mạng, cải thiện ngữ cảnh có sẵn cho thuật toán (ví dụ: biết những từ nào ngay sau và trước một từ trong câu)(1) .

(1) BiLSTM Explained | Papers With Code. <https://paperswithcode.com/method/bilstm>.

## 5.2. Xây dựng mô hình

Chúng tôi đã sử dụng hàm train\_test\_split của thư viện sklearn để chia bộ dữ liệu thành 2 tập dữ liệu train và tập dữ liệu test với tỉ lệ là 9:1. Sau đó, chúng tôi đã tiến hành training hai mô hình máy học là SVM và CNN để phân loại dữ liệu.

Mô hình SVM của chúng tôi sử dụng thư viện SVM của SKLearn để tiến hành training. Như đã trình bày ở phần trước, SVM là một thuật toán phân loại được sử dụng phổ biến trong machine learning và đã được chứng minh là có hiệu quả trong việc phân loại các bộ dữ liệu phức tạp.

Mô hình CNN của chúng tôi bao gồm các layer như sau:

* Layer Embedding là một phần quan trọng trong quá trình xử lý dữ liệu văn bản. Layer này sẽ đưa các từ trong văn bản vào một không gian vector từ với kích thước thấp hơn nhiều, giúp cho việc tính toán và phân tích các từ được dễ dàng hơn.
* Sau khi qua layer Embedding, dữ liệu được đưa vào layer Convolution1D và MaxPool1D. Layer Convolution1D sẽ tìm ra các đặc trưng và cụm từ quan trọng trong câu, từ đó giúp giảm kích thước của bộ sưu tập. Layer MaxPool1D sẽ lấy giá trị lớn nhất của các đặc trưng đã tìm được bởi layer Convolution1D, giúp tăng tính đơn giản và hiệu quả của mô hình.
* Sau đó, dữ liệu được đưa vào Layer sử dụng thuật toán LSTM/BiLSTM/Parallel LSTM-CNN/Sequential LSTM-CNN/Sequential BiLSTM-CNN. Layer này giúp mô hình có khả năng ghi nhớ thông tin trong quá khứ và sử dụng nó để dự đoán thông tin trong tương lai.
* Tiếp theo, dữ liệu sẽ được đưa qua Layer Dropout với tỉ lệ 0.2 để loại bỏ một số nơ-ron ngẫu nhiên trong quá trình huấn luyện, giúp tránh tình trạng quá khớp (overfitting) của mô hình.
* Sau đó, dữ liệu được đưa vào Layer Dense với kích thước vector đầu ra là 25 và hàm kích hoạt là relu. Hàm ReLU sẽ loại bỏ các giá trị âm trong đầu vào và giữ lại các giá trị dương, từ đó giúp mô hình học được các quan hệ phi tuyến phức tạp tiềm ẩn trong dữ liệu.
* Dữ liệu tiếp tục được đưa qua layer Flatten để làm phẳng - giảm kích thước đầu ra thành những vector một chiều.
* Cuối cùng, dữ liệu được đưa qua layer đầu ra Dense với kích thước là 1, chứa tỉ lệ tiêu cực của dữ liệu. Layer này sử dụng hàm kích hoạt sigmoid để tính toán xác suất một mẫu dữ liệu thuộc vào một trong hai loại: tục tĩu hoặc không tục tĩu.

## 5.3. Huấn luyện mô hình

**Một số siêu tham số quan trọng:**

* Loss=‘binary\_crossentropy’: Hàm mất mát binary crossentropy được sử dụng để đánh giá độ chính xác của mô hình khi dự đoán kết quả nhị phân (0 hoặc 1). Hàm này tính toán sự khác biệt giữa giá trị thực tế và giá trị dự đoán, và càng gần 0 thì mô hình càng chính xác.
* Optimizer='adam': Thuật toán tối ưu hóa (optimization algorithm) được sử dụng để cập nhật trọng số (weights) của mô hình dựa trên gradient (đạo hàm) của hàm mất mát. Thuật toán Adam kết hợp hai kỹ thuật là momentum và adaptive learning rate để tăng tốc độ và hiệu quả của quá trình học.
* Learning\_rate là siêu tham số (hyperparameter) quyết định tốc độ học của mô hình. Tốc độ học là một số dương nhỏ quyết định bước nhảy của thuật toán tối ưu hóa khi cập nhật trọng số. Nếu tốc độ học quá lớn, thuật toán có thể bỏ qua điểm tối ưu và không hội tụ. Nếu tốc độ học quá nhỏ, thuật toán có thể mất nhiều thời gian để hội tụ hoặc rơi vào điểm tối ưu cục bộ.

Auto Tuner là một công cụ của thư viện Keras giúp tìm kiếm siêu tham số tốt nhất cho mô hình tự động, giúp giảm thời gian và công sức cho việc tìm kiếm siêu tham số tốt nhất. Chúng tôi sử dụng Auto Tuner để tìm kiếm siêu tham số tối ưu cho mô hình CNN, bao gồm kích thước embedding, số lượng lớp LSTM/BiLSTM, tỉ lệ Dropout và tỉ lệ học.

Early Stopping là một kỹ thuật giúp ngăn chặn overfitting và giảm thời gian huấn luyện mô hình. Khi huấn luyện mô hình, ta sẽ chia bộ dữ liệu thành hai tập: tập huấn luyện và tập xác thực (validation set). Early stopping sẽ theo dõi độ chính xác của mô hình trên tập xác thực và dừng lại sau một số vòng lặp nếu không có sự cải thiện đáng kể.

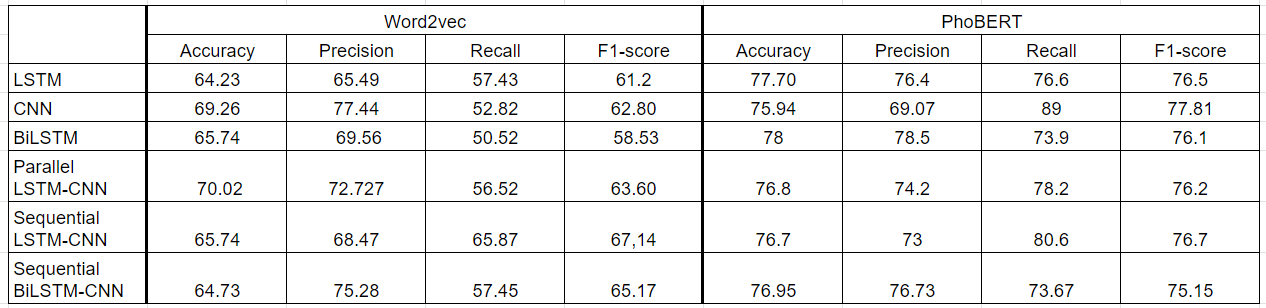
# CHƯƠNG 6: THỰC NGHIỆM VÀ ĐÁNH GIÁ KẾT QUẢ

## 6.1. Sáng tạo giai điệu

* Nhận giai điệu từ người dùng làm seed
* Phân tích giai điệu seed và trích xuất các đặc điểm key như tonic note, major/minor scale, …
* Dự đoán các nốt nhạc sẽ xuất hiện theo chuỗi thời gian
* Hiển thị các nốt lên khuông nhạc với thư viện MuseScore4
* Mã hóa các nốt nhạc và thêm vào luồng giai điệu với thư viện Music21
* Lưu luồng giai điệu dưới định dạng file MIDI
* Chuyển đổi file MIDI thành wav với FluidSynth và phát bài hát với IPython.display

## 6.2. Đánh giá kết quả

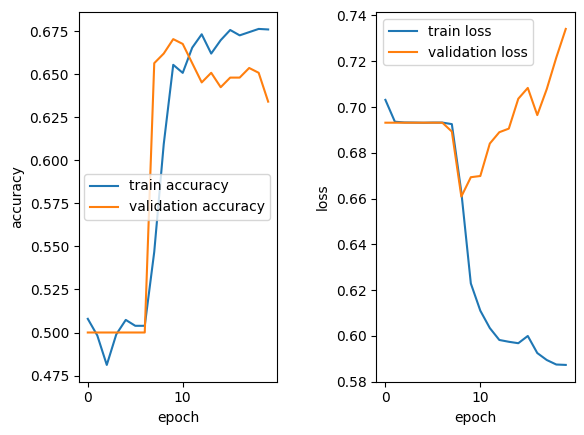
Trong nghiên cứu này, chúng tôi đã sử dụng hai phương pháp nhúng dữ liệu là Word2Vec và PhoBERT để huấn luyện và đánh giá các mô hình học máy và học sâu. Bảng dữ liệu cho thấy các kết quả tính toán của các mô hình được huấn luyện, được đo bằng các chỉ số Accuracy, Precision, Recall và F1-score. Các mô hình được so sánh trong bảng bao gồm LSTM, CNN, BiLSTM, Parallel LSTM-CNN, Sequential LSTM-CNN và Sequential BiLSTM-CNN. Kết quả cho thấy mô hình Parallel LSTM-CNN và BiLSTM có hiệu suất cao nhất đối với cả hai tập dữ liệu Word2Vec và PhoBERT, với Accuracy lần lượt là 70.02 cho Parallel LSTM-CNN và 78 cho BiLSTM.

****

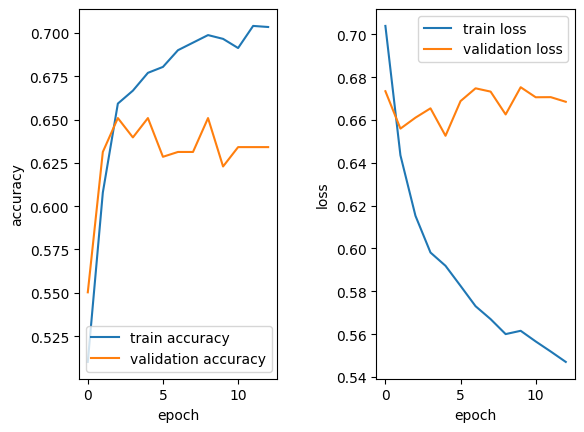
*Bảng 4.1 Kết quả so sánh giữa Word2Vec và PhoBERT*

Hình ảnh thể hiện 2 giá trị Loss và Accuracy của mô hình với số epoch được thực hiện là 100 dòng. Đường màu cam thể hiện độ chính xác của mô hình với tỷ lệ ~78%(BiLSTM của PhoBERT) và đường màu xanh biển thể hiện cho độ mất mát của mô hình. Giá trị giảm rõ rệt, tức là sai sót và lỗi được giảm, mô hình có khả năng phân loại đúng các từ ngữ có trong bộ dữ liệu.

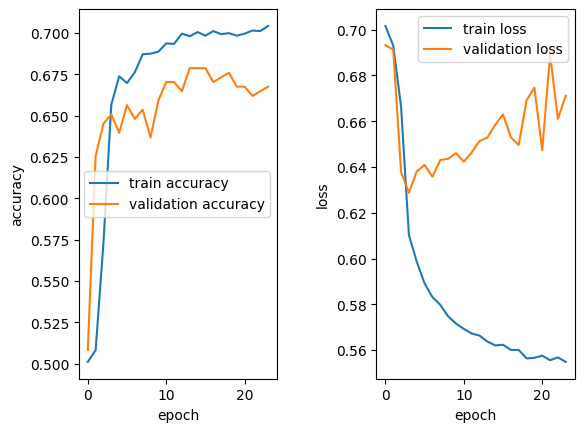
1. **Biểu đồ kết quả khi embedding bằng Word2Vec:**

****

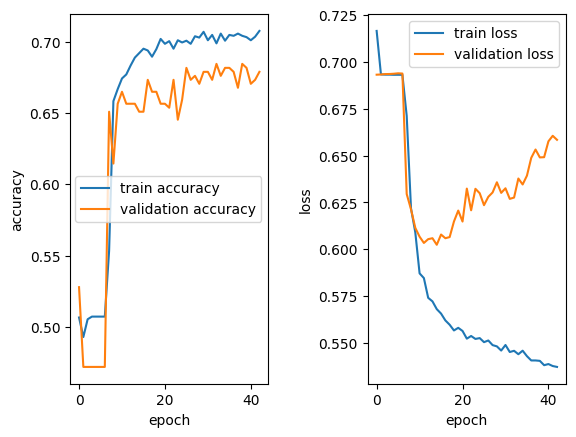
Hình 4.1 LSTM

****

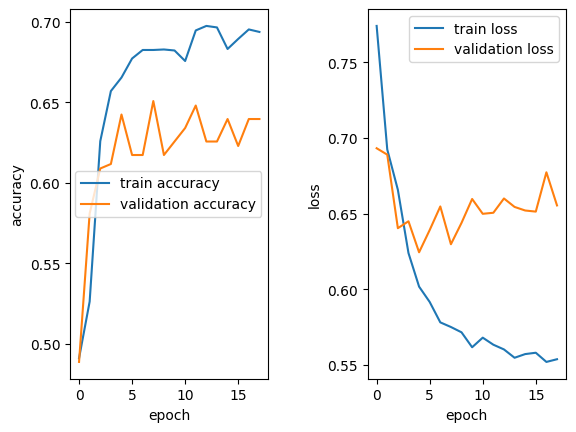
4.2 CNN

****

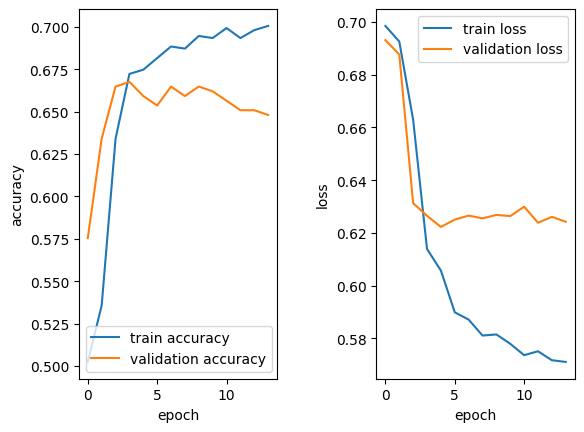
Hình 4.3 BiLSTM

****

Hình 4.4 Parallel LSTM-CNN

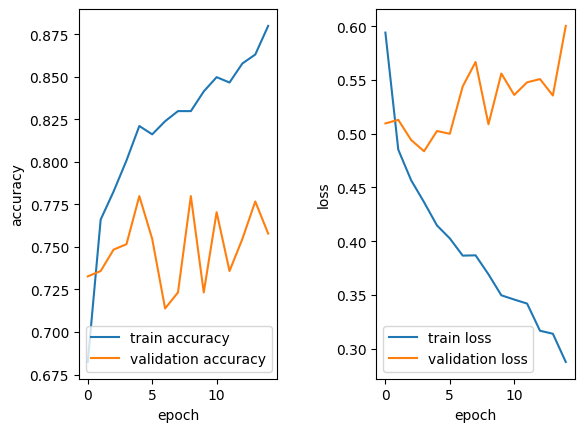
****

Hình 4.5 Sequential LSTM-CNN

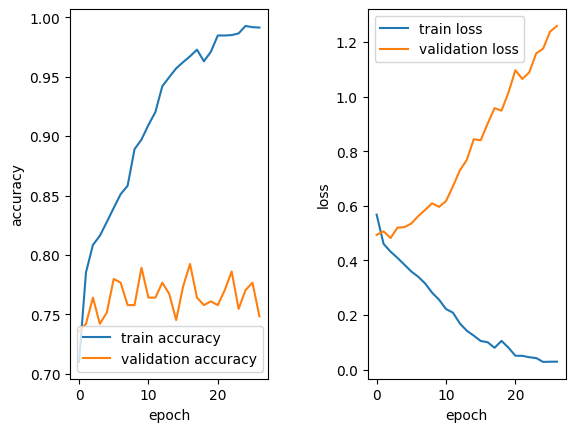
****

Hình 4.6 Sequential BiLSTM-CNN

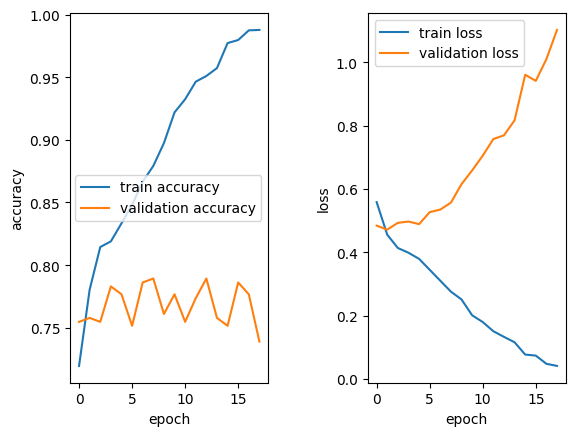
1. **Biểu đồ kết quả khi embedding bằng PhoBERT:**

****

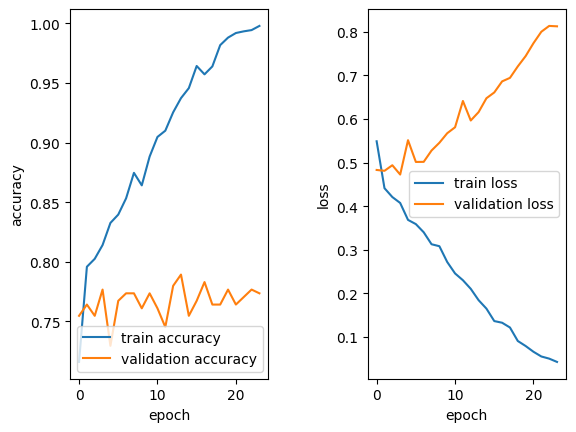
Hình 4.7 CNN



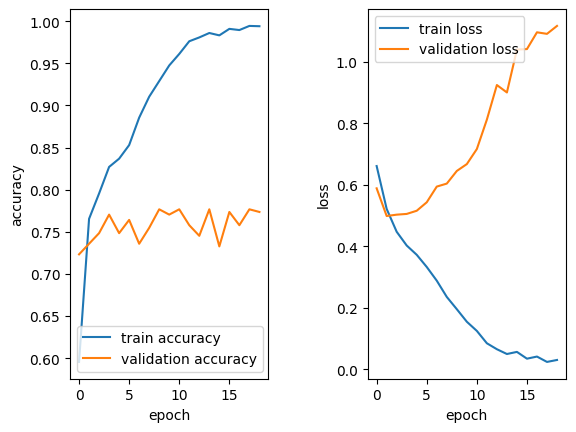
Hình 4.8 LSTM



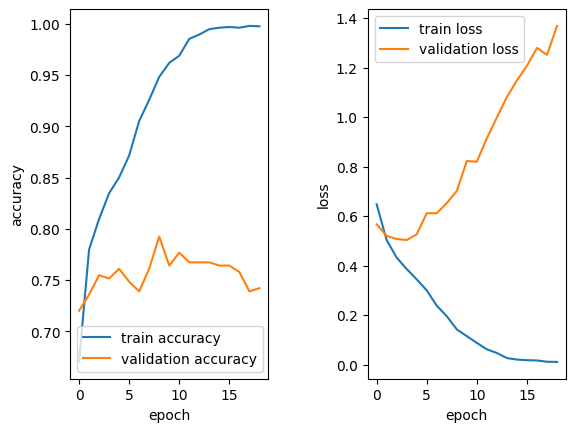
Hình 4.9 BiLSTM



Hình 4.10 Parallel LSTM-CNN



Hình 4.11 Sequential LSTM-CNN



Hình 4.12 Sequential BiLSTM-CNN

## 

Từ bảng kết quả training trên, có thể thấy mô hình PhoBERT đạt độ chính xác cao hơn so với mô hình Word2Vec trên tất cả các thang đo, bao gồm Accuracy, Precision, Recall, và F1-score. Trong đó, mô hình Parallel LSTM-CNN của PhoBERT đạt kết quả tốt nhất với độ chính xác cao nhất là 76.8% và thang đo F1-score là 76.2%. Các mô hình Sequential LSTM-CNN và BiLSTM-CNN của cả hai Word2Vec và PhoBERT đều có độ đo F1-score tương đối cao, tuy nhiên vẫn thấp hơn so với mô hình Parallel LSTM-CNN của PhoBERT. Kết quả này cho thấy PhoBERT là một giải pháp tối ưu hơn cho bài toán đang được huấn luyện so với Word2Vec.

# 

# KẾT LUẬN VÀ KIẾN NGHỊ

## 5.1 Kết luận

Trước đây, việc phát hiện nội dung thô tục trên website thường đòi hỏi phải dựa vào sự đánh giá của con người, điều này không chỉ tốn nhiều thời gian và công sức, mà còn có thể gây ra sự thiếu chính xác và mâu thuẫn trong kết quả. Bằng cách áp dụng trí tuệ nhân tạo, nghiên cứu này đã đưa ra một phương pháp mới tiềm năng để tự động phát hiện và loại bỏ nội dung thô tục trên website với độ chính xác khá cao.

Công nghệ có thể được áp dụng rộng rãi trong nhiều lĩnh vực bao gồm lĩnh vực phòng ngừa nội dung bất hợp pháp, đảm bảo an toàn cho trẻ em trên Internet, giúp quản lý nội dung trên các diễn đàn, mạng xã hội, tránh lan truyền nội dung thô tục và có hại đến người dùng nhằm góp phần tạo nên môi trường trực tuyến an toàn và lành mạnh.

## 5.2 Kiến nghị

Dựa trên kết quả nghiên cứu về việc sử dụng trí tuệ nhân tạo để phát hiện website có chứa nội dung thô tục, chúng tôi kiến nghị:

* Tập trung vào cải tiến và tăng cường độ chính xác của mô hình. Điều này có thể bao gồm việc sử dụng các kỹ thuật học sâu tiên tiến hơn kết hợp với các phương pháp khai thác ngữ cảnh và thông tin đa dạng.
* Mở rộng phạm vi để xem xét việc phát hiện các loại nội dung thô tục khác nhau trên Internet, bao gồm nội dung thô tục liên quan đến bạo lực, kỳ thị, đồi trụy, hay các loại nội dung gây hại khác.
* Xem xét các phương pháp kết hợp giữa trí tuệ nhân tạo và đánh giá của con người để đạt độ chính xác cao hơn trong các trường hợp nội dung có tính nhạy cảm và dễ bị hiểu nhầm.
* Đảm bảo tính minh bạch và công khai trong quá trình kiểm duyệt nội dung thô tục trên website. Điều này có thể bao gồm việc đề xuất các tiêu chuẩn đánh giá, công khai các phương pháp và dữ liệu huấn luyện của mô hình phát hiện nội dung thô tục, cũng như công khai các kết quả đạt được từ quá trình kiểm duyệt.
* Thúc đẩy hợp tác giữa các nhà nghiên cứu, doanh nghiệp và chính phủ để đẩy mạnh ứng dụng kết quả nghiên cứu vào thực tiễn thông qua sự hợp tác trong việc xây dựng và triển khai các giải pháp công nghệ.

# TÀI LIỆU THAM KHẢO

(1) Difficulties with developing NLP for Vietnamese. GALA Global. (n.d.). https://www.gala-global.org/knowledge-center/professional-development/articles/difficulties-developing-nlp-vietnamese

(2) Tapsai, Chalermpol & Meesad, Phayung & Haruechaiyasak, Choochart. (2016). TLS-ART: Thai Language Segmentation by Automatic Ranking Trie.

(3) Complete Guide to NLP in 2023: How It Works & Top Use Cases - AIMultiple. https://research.aimultiple.com/nlp/.

(4) NLP Tutorial - Javatpoint. https://www.javatpoint.com/nlp.

(5) Natural Language Processing (NLP): 7 Key Techniques - MonkeyLearn Blog. https://monkeylearn.com/blog/natural-language-processing-techniques/.

(6) What is Natural Language Processing? | IBM. https://www.ibm.com/topics/natural-language-processing.

(7) Top 7 Applications of NLP (Natural Language Processing). https://www.geeksforgeeks.org/top-7-applications-of-natural-language-processing/.

(8) Major Challenges of Natural Language Processing (NLP) - MonkeyLearn Blog. https://monkeylearn.com/blog/natural-language-processing-challenges/.

(9) Complete Guide to NLP in 2023: How It Works & Top Use Cases - AIMultiple. https://research.aimultiple.com/nlp/.

(10) (PDF) Using Word2Vec to process big text data - ResearchGate. https://www.researchgate.net/publication/291153115\_Using\_Word2Vec\_to\_process\_big\_text\_data.

(11) Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781.

(12) Yu, Meng-Lin & Tsai, Meng-Han. (2021). ACS: Construction Data Auto-Correction System—Taiwan Public Construction Data Example. Sustainability. 13. 362. 10.3390/su13010362.

(13)Hiểu hơn về BERT: Bước nhảy lớn của Google. (2023).https://viblo.asia/p/hieu-hon-ve-bert-buoc-nhay-lon-cua-google-eW65GANOZDO#\_kien-truc-2

(14) Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Retrieved 30 April 2023, from https://arxiv.org/abs/1810.04805

(15) Nguyen, D. Q., Nguyen, D. Q., Vu, T. X., Nguyen, T. D., & Le, Q. V. (2020). PhoBERT: Pre-trained language models for Vietnamese. arXiv preprint arXiv:2003.00744.

(16) [SVM] Thuật toán hỗ trợ máy vec-tơ Support vector machines. (2022).https://luanvanhay.org/lam-thue/svm-thuat-toan-ho-tro-may-vec-to-support-vector-machines/

(17)Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20(3), 273-297.

[1] Cen, Junjie & Li, Yongbo. (2022). A Rumor Detection Method from Social Network Based on Deep Learning in Big Data Environment. Computational Intelligence and Neuroscience. 2022. 10.1155/2022/1354233.

[2] Qian, Yunfeng & Sufian, Muhammad & Hakamy, Ahmad & Deifalla, Ahmed & Elsaid, Amr. (2023). Application of machine learning algorithms to evaluate the influence of various parameters on the flexural strength of ultra-high-performance concrete. Frontiers in Materials. 9. 1114510. 10.3389/fmats.2022.1114510.

[3] Sharma, Dibya & Dutta, Sachin & Bora, Dibya. (2020). REGA: Real-Time Emotion, Gender, Age Detection Using CNN—A Review. 115-118. 10.15439/2020KM18.

[4]

[5] Quoc Tran, K., Trong Nguyen, A., Hoang, P.G. et al. (2023). Vietnamese hate and offensive detection using PhoBERT-CNN and social media streaming data. Neural Comput & Applic 35, 573–594. https://doi.org/10.1007/s00521-022-07745-w

[6] Các tài liệu tham khảo của thầy Đặng Đại Thọ

[7] https://machinelearningcoban.com

[8] https://blog.vietnamlab.vn/