TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**TRẦN VĂN THẢO - 52000805**

**NGUYỄN VÕ CÔNG HUY - 52000765**

**XÂY DỰNG HỆ THỐNG TƯ VẤN TUYỂN SINH CHO SINH VIÊN**

**DỰ ÁN CÔNG NGHỆ THÔNG TIN**

**TÍNH TOÁN THÔNG MINH**

**VÀ HỆ THỐNG THÔNG TIN**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2024**

TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**TRẦN VĂN THẢO - 52000805**

**NGUYỄN VÕ CÔNG HUY - 52000765**

**XÂY DỰNG HỆ THỐNG TƯ VẤN TUYỂN SINH CHO SINH VIÊN**

**DỰ ÁN CÔNG NGHỆ THÔNG TIN**

**TÍNH TOÁN THÔNG MINH**

**VÀ HỆ THỐNG THÔNG TIN**

Người hướng dẫn

**TS. TRẦN LƯƠNG QUỐC ĐẠI**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2024**

**LỜI CẢM ƠN**

Lời nói đầu tiên, chúng em xin gửi lời cảm ơn chân thành nhất, đặc biệt nhất đến ***TS Trần Lương Quốc Đại*** đã hỗ trợ và truyền đạt những kiến thức quý báo cũng như đã tạo mọi điều kiện thuận lợi nhất để chúng em được tiếp cận và hoàn thành môn Dự án công nghệ thông tin.

Với thời gian và trình độ còn hạn chế, bài báo cáo không thể tránh khỏi những thiếu sót. Kính mong thầy sẽ chỉ bảo cũng như là đóng góp ý kiến để bài báo cáo của chúng em ngày càng dần hoàn thiện hơn. Đó sẽ là hành trang quý giá để chúng em có thể hoàn thiện về kĩ năng và kiến thức liên quan đến những vấn đề sau này.

Lời cuối cùng, chúng em xin kính chúc thầy thật nhiều sức khỏe, thành công, hạnh phúc và luôn giữ mãi sự nhiệt huyết để có thể giúp thêm thật nhiều thế hệ sinh viên Trường Đại Học Tôn Đức Thắng có nhiều sự tự tin và vững kiến thức với môn Dự án công nghệ thông tin. Ngọn lửa của môn học sẽ ngày càng được lan tỏa rộng rãi nhiều hơn!

*TP. Hồ Chí Minh, ngày 03 tháng 03 năm 2024*

*Tác giả*

*(Ký tên và ghi rõ họ tên)*

*N.V.C Huy*

*Nguyễn Võ Công Huy*

*T.V Thảo*

*Trần Văn Thảo*

**CÔNG TRÌNH ĐƯỢC HOÀN THÀNH**

**TẠI TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

Tôi xin cam đoan đây là công trình nghiên cứu của riêng tôi và được sự hướng dẫn khoa học của TS. Trần Lương Quốc Đại. Các nội dung nghiên cứu, kết quả trong đề tài này là trung thực và chưa công bố dưới bất kỳ hình thức nào trước đây. Những số liệu trong các bảng biểu phục vụ cho việc phân tích, nhận xét, đánh giá được chính tác giả thu thập từ các nguồn khác nhau có ghi rõ trong phần tài liệu tham khảo.

Ngoài ra, trong Dự án còn sử dụng một số nhận xét, đánh giá cũng như số liệu của các tác giả khác, cơ quan tổ chức khác đều có trích dẫn và chú thích nguồn gốc.

**Nếu phát hiện có bất kỳ sự gian lận nào tôi xin hoàn toàn chịu trách nhiệm về nội dung Dự án của mình**. Trường Đại học Tôn Đức Thắng không liên quan đến những vi phạm tác quyền, bản quyền do tôi gây ra trong quá trình thực hiện (nếu có).

*TP. Hồ Chí Minh, ngày 03 tháng 03 năm 2024*

*Tác giả*

*(Ký tên và ghi rõ họ tên)*

*N.V.C Huy*

*Nguyễn Võ Công Huy*

*T.V Thảo*

*Trần Văn Thảo*

**XÂY DỰNG HỆ THỐNG TƯ VẤN TUYỂN SINH CHO SINH VIÊN**

**TÓM TẮT**

Hệ thống tư vấn tuyển sinh sinh viên là một đề tài nghiên cứu nhằm tạo ra một hệ thống hỗ trợ tư vấn, giải đáp các thắc mắc cho các thí sinh khi xét tuyển vào các trường đại học ở Việt Nam. Đề tài này đã xây dựng một website hoàn chỉnh dựa trên MERN Stack, cung cấp thông tin tuyển sinh, các khóa học cho các thí sinh có mong muốn học trước ngành Công nghệ thông tin và hỗ trợ tư vấn cho các thí sinh trong quá trình xét tuyển. Hệ thống này giúp tối ưu hóa việc tư vấn, nâng cao chất lượng đào tạo, và hạn chế tình trạng bỏ học giúp các thí sinh có thể đưa ra quyết định chính xác nhất cho tương lai của mình.

Ngoài ra, xu hướng sử dụng chatbot đang được ưa chuộng trên thế giới. Việc áp dụng chatbot vào tư vấn tuyển sinh giúp tối ưu hiệu quả vận hành, tóm tắt và cung cấp thông tin nhanh chóng, đầy đủ nhất cho thí sinh.

**STUDENT ADMISSION ADVISORY SYSTEM**

**ABSTRACT**

Student Admission Advisory System is a research topic that aims to create a system of advisory support, answering questions for applicants when applying to universities in Vietnam. This topic has built a complete website based on MERN Stack, providing admission information, courses for candidates who wish to pre-study in Information Technology and advising candidates during the admissions process. This system optimizes counselling, improves the quality of training, and limits dropouts, enabling candidates to make the most accurate decisions for their future.

Besides, the trend of using chatbots is popular around the world. The use of chatbots in the admissions consultancy helps to optimize operational efficiency, summarize and provide the candidates with the fastest, most complete information.

**MỤC LỤC**

[DANH MỤC HÌNH VẼ vii](#_Toc159858300)

[DANH MỤC BẢNG BIỂU viii](#_Toc159858301)

[DANH MỤC CÁC CHỮ VIẾT TẮT ix](#_Toc159858302)

[CHƯƠNG 1. MỞ ĐẦU VÀ TỔNG QUAN ĐỀ TÀI 1](#_Toc159858303)

[1.1 Lý do chọn đề tài 1](#_Toc159858304)

[1.2 Mục tiêu thực hiện đề tài 1](#_Toc159858305)

[CHƯƠNG 2. CƠ SỞ LÝ THUYẾT 2](#_Toc159858306)

[2.1 Module Chatbot 2](#_Toc159858307)

[2.2 Module Website 2](#_Toc159858308)

[2.2.1 a 2](#_Toc159858309)

[2.2.2 b 2](#_Toc159858310)

[2.2.3 c 2](#_Toc159858311)

[CHƯƠNG 3. MÔ HÌNH ĐỀ XUẤT 2](#_Toc159858312)

[3.1 Mô hình classicIR(Retriever) 2](#_Toc159858313)

[3.1.1 Giới thiệu mô hình classicIR 2](#_Toc159858314)

[3.1.2 Phương pháp tìm kiếm thông tin 2](#_Toc159858315)

[3.2 Mô hình BiLSTM(Reader) 2](#_Toc159858316)

[3.2.1 Giới thiệu mô hình BiLSTM 2](#_Toc159858317)

[3.2.2 Quá trình đọc và trích xuất thông tin 2](#_Toc159858318)

[3.3 Mô hình PhoBERT (Generative) 2](#_Toc159858319)

[3.3.1 Giới thiệu mô hình PhoBERT 2](#_Toc159858320)

[3.3.2 Tạo câu trả lời tự nhiên 2](#_Toc159858321)

[CHƯƠNG 4. THỰC NGHIỆM 3](#_Toc159858322)

[4.1 Dữ liệu thực nghiệm 3](#_Toc159858323)

[4.2 Cài đặt thực nghiệm 3](#_Toc159858324)

[CHƯƠNG 5. KẾT QUẢ ĐẠT ĐƯỢC 3](#_Toc159858325)

[5.1 Mô hình <tên mô hình> của Module ChatBot 3](#_Toc159858326)

[5.2 Front-end của Module ChatBot 4](#_Toc159858327)

[5.3 Front-end của Module Website 4](#_Toc159858328)

[CHƯƠNG 6. KẾT LUẬN 5](#_Toc159858329)

[6.1 Kết luận 5](#_Toc159858330)

[6.2 Hướng phát triển 5](#_Toc159858331)

[TÀI LIỆU THAM KHẢO 6](#_Toc159858332)

# DANH MỤC HÌNH VẼ

**No table of figures entries found.**

# DANH MỤC BẢNG BIỂU

[Bảng 4.1 : Thống kê kiểu thực thể trong tập VLSP 2016 4](#_Toc1673580053)

# DANH MỤC CÁC CHỮ VIẾT TẮT

|  |  |
| --- | --- |
| BERT | Bidirectional Encoder Representations from Transformers |
| GEC | Grammatical Error Correction |
| MLM | Masked Language Model |
| NLP | Natural Language Processing |
| NSP | Next Sentence Prediction |

# MỞ ĐẦU VÀ TỔNG QUAN ĐỀ TÀI

## Lý do chọn đề tài

Một trong những lý do quan trọng để chọn đề tài này là đáp ứng nhu cầu ngày càng tăng về thông tin trong quá trình tuyển sinh. Hệ thống tuyển sinh truyền thống thường gặp khó khăn trong việc cung cấp thông tin nhanh chóng và chính xác đối với các học sinh đang tìm hiểu về quy trình tuyển sinh. Sự chậm trễ trong phản hồi có thể tạo ra rủi ro về việc học sinh không nhận được thông tin đầy đủ và kịp thời, ảnh hưởng đến quyết định của họ trong việc chọn lựa trường học.

Hệ thống Q&A có khả năng ngay lập tức cung cấp thông tin chi tiết và rõ ràng, giúp học sinh giải đáp mọi thắc mắc một cách hiệu quả. Bằng cách này, đề tài không chỉ giải quyết vấn đề bất cập trong quá trình tuyển sinh mà còn đáp ứng nhu cầu ngày càng cao của học sinh đối với thông tin nhanh chóng và dễ tiếp cận.

## Mục tiêu thực hiện đề tài

Mục tiêu chính của nghiên cứu là phát triển một hệ thống Q&A trong quá trình tuyển sinh, nhằm giảm thời gian phản hồi, đa dạng hóa khả năng trả lời, tự động cập nhật thông tin, và thực nghiệm ứng dụng trong môi trường thực tế. Mục tiêu này nhằm tối ưu hóa trải nghiệm thông tin cho học sinh, giảm công sức của tổ chức tuyển sinh, và tạo cơ sở cho phát triển tương lai trong ứng dụng trí tuệ nhân tạo trong giáo dục.

# CƠ SỞ LÝ THUYẾT

## Module Chatbot

## Module Website

### MongoDB và Mongoose

#### Tổng quan về MongoDB:

…

#### Ưu và nhược điểm của MongoDB:

…

#### Tổng quan về Mongoose:

…

#### Ưu và nhược điểm của Mongoose:

…

### ExpressJS

#### Tổng quan:

…

#### Ưu và nhược điểm:

…

### ReactJS

#### Tổng quan:

…

#### Ưu và nhược điểm:

…

### NodeJS

#### Tổng quan:

…

#### Ưu và nhược điểm:

…

# MÔ HÌNH ĐỀ XUẤT

## Mô hình classicIR(Retriever)

### Giới thiệu mô hình classicIR

### Phương pháp tìm kiếm thông tin

## Mô hình BiLSTM(Reader)

### Giới thiệu mô hình BiLSTM

### Quá trình đọc và trích xuất thông tin

## Mô hình PhoBERT (Generative)

### Giới thiệu mô hình PhoBERT

### Tạo câu trả lời tự nhiên

# PHÂN TÍCH HỆ THỐNG

## Module Chatbot

…

## Module Website

### Các tác nhân của hệ thống

…

### Các yêu cầu chức năng

…

### Các yêu cầu phi chức năng

...

### Use cases

…

### Lược đồ quan hệ thực thể - ERD

…

# THỰC NGHIỆM

## Dữ liệu thực nghiệm

Dữ liệu trong VLSP-2016 đã được chia sẵn thành hai tập huấn luyện và kiểm thử. Thống kê chi tiết của các kiểu thực thể và các tập dữ liệu được mô tả trong Bảng 4.1**.**

Bảng 4.1: Thống kê kiểu thực thể trong tập VLSP 2016

|  |  |  |
| --- | --- | --- |
| **Label** | **Train** | **Test** |
| **LOC** | 1210 | 1377 |
| **ORG** | 7478 | 274 |
| **PER** | 6230 | 1294 |
| **MISC** | 250 | 47 |
| **Total** | 15168 | 2992 |

## Cài đặt thực nghiệm

…

# KẾT QUẢ ĐẠT ĐƯỢC

## Mô hình <tên mô hình> của Module ChatBot

…

## Giao diện của Module ChatBot

…

## Giao diện của Module Website

# KẾT LUẬN

## Kết luận

…

## Hướng phát triển

….

# TÀI LIỆU THAM KHẢO

Tiếng Việt

…

Tiếng Anh

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-term Memory. *Neural Computation*, *9*, 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2023). *Attention Is All You Need* (arXiv:1706.03762). arXiv. https://doi.org/10.48550/arXiv.1706.03762