

TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM
TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG



Công trình Nghiên cứu khoa học sinh viên năm học 2023 - 2024

XÂY DỰNG CHATBOT CHO HỆ THỐNG TƯ VẤN TÂM LÝ SINH VIÊN ĐẠI HỌC TÔN ĐỨC THẮNG

ĐƠN VỊ KHOA CÔNG NGHỆ THÔNG TIN

GIẢNG VIÊN HƯỚNG DẪN:

PGS.TS. LÊ ANH CƯỜNG

SINH VIÊN/NHÓM SINH VIÊN THỰC HIỆN:

- 1. NGUYỄN MINH CHÍ*
- 2. VẦN THỊ THANH HƯƠNG*
- 3. DƯƠNG TRỌNG CHÍ*

TP. Hồ Chí Minh, tháng 05 năm 2024

**TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM
TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG
KHOA CÔNG NGHỆ THÔNG TIN**



**CÔNG TRÌNH NGHIÊN CỨU KHOA HỌC SINH VIÊN NĂM
HỌC 2023-2024**

**XÂY DỰNG CHATBOT CHO HỆ
THỐNG TƯ VẤN TÂM LÝ SINH VIÊN
ĐẠI HỌC TÔN ĐỨC THẮNG**

Người hướng dẫn: **PGS.TS. LÊ ANH CƯỜNG**

Người thực hiện: **NGUYỄN MINH CHÍ - 52000189**

VĂN THỊ THANH HƯƠNG – 52200128

DƯƠNG TRỌNG CHÍ - 52000742

THÀNH PHỐ HỒ CHÍ MINH, NĂM 2024

LỜI CẢM ƠN

Lời đầu tiên, chúng em xin chân thành gửi lời cảm ơn chân đến *Ban giám hiệu trường Đại Học Tôn Đức Thắng* đã tạo điều kiện cho chúng em hoàn thành công trình nghiên cứu này thông qua hệ thống thư viện đa dạng tài liệu hay, bổ ích.

Chúng em xin chân thành cảm ơn **Thầy Lê Anh Cường**, đã giảng dạy và truyền đạt kiến thức một cách tận tình và chi tiết, cung cấp tài liệu tham khảo giúp chúng em đủ nền tảng để vận dụng vào việc viết bài báo cáo này.

Trong quá trình làm bài báo cáo nghiên cứu, chúng em khó tránh khỏi thiếu sót là điều chắc chắn, chúng em rất mong nhận được những ý kiến đóng góp quý báu của quý thầy cô để kiến thức của chúng em được hoàn thiện hơn.

Sau cùng em xin chân thành cảm ơn và kính chúc quý thầy cô trong khoa Công Nghệ Thông Tin luôn khỏe mạnh và thành công trong sự nghiệp giảng dạy.

LỜI CAM KẾT

Tôi xin cam đoan đây là sản phẩm nghiên cứu của riêng chúng tôi và được sự hướng dẫn của PGS.TS. Lê Anh Cường. 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 nghiên cứu 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 nghiên cứu 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 tháng năm

Tác giả

(ký tên và ghi rõ họ tên)

Nguyễn Minh Chí

Dương Trọng Chí

Văn Thị Thanh Hương

MỤC LỤC

MỤC LỤC.....	3
TÓM TẮT	1
CHƯƠNG 1 – MỞ ĐẦU	2
1.1 Đặt vấn đề	2
1.2 Lý do chọn đề tài.....	3
1.3 Đối tượng nghiên cứu	5
1.4 Phạm vi nghiên cứu.....	5
1.5 Những mô hình AI hiện nay	5
1.6 Những vấn đề hạn chế trong phát triển AI.....	7
1.7 Phương pháp giải quyết	9
CHƯƠNG 2 – CƠ SỞ LÝ THUYẾT	10
2.1. Mô hình ngôn ngữ lớn.....	10
2.1.1 Llama 3	10
2.1.2 Vistral và Vinallama	11
2.2 QLoRA	12
2.2.1 Khó khăn với mô hình 7 tỷ tham số.....	12
2.2.2 QLoRA is all you need	13
CHƯƠNG 3 – PHƯƠNG PHÁP ĐỀ XUẤT	14
3.1 Pre-trained các mô hình ngôn ngữ lớn.....	14
3.1.1. Lựa chọn mô hình ngôn ngữ lớn.....	14
3.1.2. Mục tiêu của việc pre-training	14
3.1.3. Kỹ thuật pre-training	15
3.2 Instruction Fine-tune cho bài toán chatbot.....	15
3.2.1. Lý do lựa chọn kỹ thuật hướng dẫn	15
3.2.2. Kỹ thuật instruction fine-tuning.....	15
3.2.3. Áp dụng kỹ thuật instruction fine-tuning cho các mô hình	16
3.3 Lượng tử hóa (quantization) mô hình ngôn ngữ lớn.....	16
3.3.1. Cần lượng tử hóa mô hình	16

3.3.2. Lượng tử hóa là gì và mục đích	16
3.3.3. Lợi ích của lượng tử hóa	17
3.3.4. Áp dụng lượng tử hóa cho các mô hình	17
3.4 Dữ liệu	18
3.4.1. Nguồn dữ liệu	18
3.4.2. Tiền xử lý dữ liệu	18
3.4.3. Các loại dữ liệu	18
CHƯƠNG 4: KẾT QUẢ THỰC NGHIỆM	20
4.1 Dữ liệu	20
4.1.1 Dữ liệu cho text completion fine-tuning	20
4.1.2 Dữ liệu tham vấn tâm lý	20
4.2 Fine-tune	21
4.2.1 Text completion fine-tuning	21
4.2.2 Instruction fine-tuning	22
4.2.3 Kết quả đạt được	23
4.3 Xây dựng chatbot	26
4.3.1 Ollama	26
4.3.2 Streamlit	26
4.4 Ý nghĩa kết quả nghiên cứu	28
4.4.1 Ý nghĩa khoa học	28
4.4.2 Ý nghĩa thực tiễn	28
4.4 Đóng góp của nghiên cứu	28
4.5 Những điều đạt được	29
4.6 Định hướng nghiên cứu tiếp theo	29
TÀI LIỆU THAM KHẢO	30

TÓM TẮT

Đề tài nghiên cứu khoa học này với mục tiêu hiểu tâm lý học sinh sinh viên đang gặp phải như trầm cảm, lo âu cũng như tìm ra cách giải quyết tư vấn các vấn đề thông qua hình thức AI, chatbot, có thể vừa phù hợp thời đại công nghệ 4.0, vừa hỗ trợ kịp thời giúp người dùng giảm bớt các triệu chứng và có thể thư giãn hơn.

Đề tài nghiên cứu gồm bốn chương:

- Chương 1: Mở đầu – Nội dung gồm đặt vấn đề, nêu lý do chọn đề tài, sự cần thiết của đề tài, phạm vi, đối tượng nghiên cứu, tóm lược các giải pháp, nêu những các AI liên quan đã có và giải quyết các vấn đề còn tồn tại của các phương án nghiên cứu trước đó.
- Chương 2: Cơ sở lý thuyết – Nội dung gồm các cơ sở lý thuyết, mục tiêu công trình.
- Chương 3: Phương pháp đề xuất – Nêu phương pháp giải quyết vấn đề
- Chương 4: Kết quả và định hướng tương lai – Nội dung gồm các kết quả nghiên cứu đạt được, kết luận đề tài, ý nghĩa nghiên cứu khoa học, và hiệu quả của nghiên cứu đối với các mặt đời sống, định hướng phát triển mở rộng phạm vi nghiên cứu trong tương lai.

CHƯƠNG 1 – MỞ ĐẦU

1.1 Đặt vấn đề

Đất nước đang càng ngày càng phát triển nên đòi hỏi phải học tập không ngừng, đặc biệt là các thế hệ trẻ hiện nay. Chính vì thế việc học đã trở nên áp lực hơn, đè nặng hơn, đôi mặt với nhiều kì thi cử, định hướng tương lai và đặc biệt là vấn đề liên quan đến môi trường xã hội xung quanh, không ngừng thay đổi và biến động. Điều này dẫn đến các vấn đề tâm lý như lo âu, trầm cảm, căng thẳng lên những thế hệ trẻ, đặc biệt là những sinh viên đại học ngày nay. Hơn nữa, những sinh viên Đại học đa số đều vừa đi học, vừa đi làm với mong muốn đỡ được gánh nặng đồng tiền ba mẹ cũng như chăm lo kinh phí cho bản thân.

Đại học Tôn Đức Thắng là đại học thuộc top hàng đầu Việt Nam. Trường có những cơ sở vật chất hiện đại và chương trình học chuẩn hóa quốc tế. Tuy nhiên, điều này cũng đồng nghĩa với việc các chương trình học mỗi môn đều có những độ khó cao và đòi hỏi những kiến thức chuyên sâu nền tảng hơn. Dẫn đến việc học trở nên quan trọng, việc thi cử cũng đặt ở hàng đầu. Thêm vào đó, áp lực đồng trang lứa chính là những áp lực khá phổ biến tại trường, khiến cho nhiều bạn sinh viên cảm thấy mình phải nỗ lực hơn để không thua kém bạn bè, vì có nhiều bạn có những thành tích xuất sắc, tham gia nhiều hoạt động sự kiện.

Theo như khảo sát trên diễn đàn sinh viên Đại học Tôn Đức Thắng, thì đa phần các bạn sinh viên ngày nay đều có trên tay một chiếc điện thoại thông minh và sử dụng nó ít nhất 4 tiếng/ngày, thậm chí con số tăng lên cao hơn cả thế. Việc này sẽ hạn chế sự tiếp xúc của các bạn với mọi thứ xung quanh. Thậm chí có nhiều bạn từ chối việc phản ứng với môi trường xung quanh, khiến cho bản thân sinh ra cảm xúc e ngại giao tiếp, đặc biệt là các vấn đề riêng tư về cá nhân. Nên có thể nhiều bạn có nguy cơ mắc những chứng bệnh về tâm lý đều sẽ che giấu và tỏ ra không có gì, làm cho tâm lý bản thân trở nên trầm trọng hơn. Đặc biệt ở lứa tuổi 18, 20 này đều che giấu cảm xúc, để mọi thứ trong lòng và không muốn ai biết điều mình gặp phải.

Nhận thấy những vấn đề này, trường Đại học Tôn Đức Thắng đã thành lập và vận hành Phòng tư vấn tâm lý học đường – Phòng E001. Những bạn sinh viên có thể

chọn lịch và đến ngày hẹn để đi tư vấn. Ngoài ra, trường cũng tổ chức các buổi chuyên đề tâm lý, giảm stress trong việc học hành cho các bạn sinh viên tham gia, nhằm hỗ trợ các bạn giảm đi các cảm xúc tiêu cực trong mình. Tuy trường có những giải pháp phục vụ cho tâm lý sinh viên nhưng vẫn có những hạn chế như sau:

- Các bạn muốn tư vấn đều phải đặt lịch hẹn trước trên cổng thông tin sinh viên, nhưng số thời gian và những bạn sinh viên muốn tham gia tư vấn trong 1 ngày đều có giới hạn số người. Nhưng lịch hẹn chỉ hiển thị trước 1 ngày. Chính vì điều đó mà có nhiều bạn vẫn chưa thể đặt lịch hẹn vì hết thời gian trống, và các bạn phải đợi những ngày khác. Nếu như chờ đợi quá lâu sẽ làm cho bản thân không muốn đi nữa, như thế nhiều cảm xúc tiêu cực tích tụ nhiều ngày dần.
- Các buổi chuyên đề, workshop muốn hiệu quả đều giới hạn sinh viên, tầm khoảng 10-20 người. Tuy một tuần có nhiều buổi tổ chức nhưng sinh viên thì rất nhiều, chính vì thế những bạn không tham gia được ngay lúc bản thân cần.

Hơn thế nữa, chúng ta làm thế nào để ngăn chặn, phát hiện sớm những dấu hiệu có tâm lý bất thường mà đối tượng là các bạn sinh viên của Tôn Đức Thắng tự nguyện chia sẻ vấn đề mình gặp phải? Liệu có giải pháp nào tối ưu, ngay tức thời, ngay lúc sinh viên cần nhất nhưng không cần phải chờ đợi, mà phù hợp với lứa tuổi sinh viên, giúp cho các bạn sinh viên mở lòng hơn, chia sẻ những khúc mắc của bản thân và có những suy nghĩ đúng đắn hơn về bản thân cũng như nhận thức được vấn đề của bản thân?

1.2 Lý do chọn đề tài

Tâm lý là một khía cạnh đóng vai trò then chốt trong quá trình học tập và sự phát triển cá nhân của con người hiện đại nói chung và của sinh viên đại học Tôn Đức Thắng nói riêng. Sự tiến bộ không ngừng của công nghệ và xã hội hiện đại đặt ra một loạt các thách thức mới cho tâm lý con người. Mặc dù chúng ta đã thấy những bước tiến trong lĩnh vực công nghệ, não bộ con người vẫn duy trì cấu trúc cơ bản từ thời tiền sử. Sự không cân đối này đã tạo ra tình huống khi chúng ta đôi khi gặp nhiều vấn

đề tâm lý trong cuộc sống hàng ngày, và để đối phó với chúng, tư vấn tâm lý đã trở thành một yếu tố quan trọng và thiết yếu trong xã hội ngày nay, đặc biệt là thế hệ sinh viên đại học hiện nay, khi họ ở trong độ tuổi nằm giữa sự trẻ con và trưởng thành, và đặc biệt nhạy cảm với nhiều vấn đề trong cuộc sống.

Có thể thấy, tư vấn tâm lý là một vấn đề hết sức quan trọng đối với sinh viên. Đầu tiên, tâm lý sinh viên có thể bị ảnh hưởng bởi học tập và các kỳ thi quan trọng. Sự lo lắng, căng thẳng và áp lực có thể ảnh hưởng đến hiệu suất học tập và sức khỏe tinh thần của họ. Tư vấn tâm lý giúp sinh viên xây dựng cách ứng phó với áp lực này, tạo ra một tâm lý tích cực và khám phá tiềm năng cá nhân. Thứ hai, trong thời kỳ đại học, nhiều sinh viên trải qua sự thay đổi lớn về cuộc sống và vai trò xã hội. Họ có thể gặp khó khăn trong việc xác định bản thân, xây dựng mối quan hệ xã hội, và đối diện với những thách thức trong cuộc sống cá nhân. Tư vấn tâm lý giúp họ hiểu rõ hơn về bản thân, giải quyết xung đột và thúc đẩy sự phát triển xã hội và cá nhân. Ngoài ra, tư vấn tâm lý còn có vai trò quan trọng trong việc đảm bảo sự việc phát triển sự cân bằng tinh thần và sức khỏe tinh thần cho sinh viên. Nó có thể giúp phát hiện và điều trị các vấn đề tâm lý như căng thẳng, lo âu, và trầm cảm sớm, từ đó ngăn chặn sự trầm trọng hóa của các vấn đề này.

Mặc dù hiện nay có nhiều chatbot thông minh như Chat GPT, Bard và nhiều loại khác, thế nhưng, chưa có một chatbot đặc biệt dành riêng cho tư vấn tâm lý của sinh viên đại học Tôn Đức Thắng. Nhiều người dựa vào các chatbot tổng quát như Chat GPT để tìm hiểu về tâm lý, tuy nhiên, vì chúng chuyên về nhiều lĩnh vực khác nhau, chúng thường cung cấp những câu trả lời chung chung, không tập trung vào tâm lý cụ thể. Kết quả là, các câu trả lời có thể trở nên không thỏa đáng hoặc gây hiểu lầm.

Nếu như các bạn sinh viên có trở ngại giao tiếp giữa người với người thì giao tiếp với AI sẽ một phần nào đó giúp các bạn sinh ra tâm lý yên tâm, thoải mái, hoàn toàn tự nguyện, cảm xúc cũng được bộc lộ một cách nguyên thủy nhất. Bởi vì khi các bạn có dấu hiệu có liên quan đến một số bệnh tâm lý sẽ muốn cho bản thân vào vùng an toàn, không muốn mạo hiểm cũng như đánh cược điều gì. Chính vì thế trò chuyện

cùng AI là giải pháp hầu như các bạn sinh viên đều lựa chọn vì các bạn lúc đó cảm thấy mình được an toàn. Hơn nữa việc phát triển chatbot tư vấn tâm lý cho sinh viên Đại học Tôn Đức Thắng cũng hỗ trợ cho việc tư vấn trực tiếp tại E001, giảm bớt số lượng sinh viên, cũng như đưa ra những giải pháp kịp thời, ngay tức thì mà những bạn sinh viên Tôn Đức Thắng không cần phải chờ đợi và đặt lịch hẹn tại E001.

Hiểu được những vấn đề trên cũng như nắm bắt các xu hướng công nghệ mới trên thế giới về AI và các mô hình ngôn ngữ lớn, nhóm đề xuất ra một hệ thống chatbot chuyên cung cấp các câu trả lời tư vấn tâm lý cho các đối tượng sinh viên đại học Tôn Đức Thắng, mang lại các câu trả lời có giá trị hơn, có tính riêng của đại học Tôn Đức Thắng. Mô hình này nhằm mang lại sự hỗ trợ toàn diện, thông qua các câu trả lời đúng đắn, thúc đẩy tâm lý tích cực, và tạo ra môi trường thúc đẩy sự phát triển cá nhân của sinh viên khi học tập tại trường Tôn Đức Thắng.

1.3 Đối tượng nghiên cứu

Trong đề tài này, đối tượng nghiên cứu bao gồm:

- Các phương pháp thu thập dữ liệu về tâm lý, chủ yếu thông qua Google Form, dữ liệu từ các tình huống thực tế, blog, prompt engineering.
- Các loại bệnh về tâm lý thường gặp đối với sinh viên học tại trường đại học Tôn Đức Thắng.
- Các mô hình học máy, học sâu, mô hình ngôn ngữ lớn dùng cho bài toán chatbot.

1.4 Phạm vi nghiên cứu

Trong bài nghiên cứu này chúng tôi sẽ tập chung vào các vấn đề chính sau:

- Dự đoán các loại bệnh về tâm lý và đưa ra các giải pháp phù hợp cho từng sinh viên học tập tại trường Tôn Đức Thắng ở độ tuổi từ 18 đến 24 tuổi.
- Xây dựng chatbot tư vấn tâm lý dành cho sinh viên với ngôn ngữ chính là tiếng Việt, và mang tính đại học Tôn Đức Thắng.

1.5 Những mô hình AI hiện nay

Đề tài nghiên cứu về tâm lý là đề tài rất nổi bật và có nhiều tác giả đi trước đã nghiên cứu và kết hợp công nghệ khoa học vào nghiên cứu đề tài.

Trước đây, các chatbot về tư vấn tâm lý nói riêng và chatbot nói chung, thường là các chatbot dạng rule-based, tức là các câu trả lời được lập trình sẵn để trả lời. Trong khi đó, các vấn đề về tư vấn tâm lý thì rất đa dạng, nhiều tính huống, ngữ cảnh, hoàn cảnh, do đó, việc lập trình rule-based cho chatbot là không khả thi.

Ngay sau sự ra đời của bài báo Attention is all you need, các mô hình Transformers trở nên bùng nổ, ngày càng nhiều các mô hình Transformers có khả năng giải quyết các vấn đề phức tạp hơn. Chẳng hạn, Google cho ra mắt BERT vào năm 2018, được cho là transformer có performance tốt nhất lúc bấy giờ, hoạt động tốt trên các bài toán như sentiment analysis, . Hay mô hình T5 được xem như là một đỉnh cao công nghệ với hiệu suất cao cho các bài toán dịch máy, tóm tắt văn bản. Hoặc GPT-2 của OpenAI.

Với sự ra đời của ChatGPT vào tháng 11 năm 2023, thế giới lao vào cuộc đua Generative AI của các ông lớn công nghệ như Meta, Google, OpenAI. Các start-up về AI cũng ngày càng sôi nổi và cho ra đời các mô hình open-source chất lượng như llama2, llama3, Mistral,... Các mô hình này đạt được các kết quả cao trên các benchmark phổ biến, mở ra một kỉ nguyên mới cho Generative AI.

Tuy nhiên, dù GenerativeAI có hiện đại đến mấy, có thể làm các tác vụ có độ phức tạp cao, thì bài toán chatbot vẫn luôn là một vấn đề nóng hổi, vẫn là một mảnh đất màu mỡ cho sự phát triển của GenAI.

Bên cạnh đó, khi xã hội càng phát triển, nó làm cuộc sống của con người ngày càng tiện nghi hơn. AI được ví như một phát minh có tầm quan trọng như điện năng của thế kỷ trước. Tuy thế, khoa học cũng làm cho tâm hồn con người trở nên nghèo nàn đi. Sự ra đời của smartphone và mạng xã hội khiến cho con người dần ngại giao tiếp, ngại kết nối. Con người trở nên sống khép kín hơn. Hằng ngày, khi mở mạng xã hội, người ta sẽ bị ngập trong các tin tức tiêu cực, ngu ngốc, hay những thành tựu của bạn bè khiến họ áp lực đồng trang lứa. Điều đó gây nên một áp lực hữu hình lẫn vô hình giữa các bạn trẻ. Do áp lực từ nhiều thứ, cộng thêm việc ít người thấu hiểu và chia sẻ, tâm lý giới trẻ ngày càng trở nên bất ổn.

Mặc dù nhiều mô hình AI ngày càng ra đời, bắt kịp xu thế nhưng đa số đều không mang những tính riêng về cụ thể cái gì đó. Nếu bạn là sinh viên Tôn Đức Thắng và muốn hỏi những vấn đề mình gặp phải tại trường thì câu trả lời của nó cũng mang tính chất đại trà dành cho các sinh viên các trường. Như thế không giúp sinh viên Tôn Đức Thắng tìm ra giải pháp mà làm các bạn ngao ngán nó.

Nắm bắt xu hướng về Generative AI, LLMs, và với vấn đề tâm lý nóng hổi của sinh viên đại học Tôn Đức Thắng hiện nay, nhóm quyết định nghiên cứu đề tài chatbot Tham vấn tâm lý cho đối tượng sinh viên Tôn Đức Thắng. Với mục tiêu giúp sinh viên học tập tại trường có cái nhìn rõ hơn về các vấn đề tâm lý bản thân, cũng như được quan tâm, chia sẻ, thấu hiểu. Hơn nữa những câu trả lời đều mang tính Đại học Tôn Đức Thắng, và những giải pháp thực tế áp dụng ngay trong trường hợp bản thân. Đây là đề tài vừa mang tính nhân văn cao vừa ứng dụng được các công nghệ tiên tiến, là một sự kết hợp hoàn hảo giữa Khoa học máy tính và Khoa học xã hội, với mong muốn hỗ trợ các bạn sinh viên học tập tốt hơn tại trường Đại học Tôn Đức Thắng.

1.6 Những vấn đề hạn chế trong phát triển AI

Trong bối cảnh ứng dụng trí tuệ nhân tạo tổng hợp (GenAI) vào lĩnh vực tham vấn tâm lý, một số vấn đề còn tồn tại và cần được nghiên cứu để cải thiện hiệu quả và độ tin cậy của các chatbot. Dưới đây là một số vấn đề chính:

- Hiểu biết ngữ cảnh và cảm xúc:
 - o Vấn đề: Chatbot hiện tại vẫn gặp khó khăn trong việc hiểu đúng ngữ cảnh và cảm xúc của người dùng. Điều này dẫn đến việc trả lời không phù hợp hoặc thiếu nhạy cảm với trạng thái tâm lý của người dùng.
 - o Hướng nghiên cứu: Sử dụng các mô hình ngôn ngữ lớn (LLMs) tiên tiến như LLaMA 3 hoặc Mistral để cải thiện khả năng hiểu ngữ cảnh và cảm xúc. Các LLMs này có khả năng xử lý lượng lớn dữ liệu và học từ ngữ cảnh phức tạp, từ đó cung cấp các phản hồi chính xác và nhạy cảm hơn.
- Tính nhất quán và liên tục trong phản hồi:

- Vấn đề: Một số chatbot gặp khó khăn trong việc duy trì tính nhất quán và liên tục khi tham gia vào các cuộc trò chuyện kéo dài. Điều này có thể làm gián đoạn quá trình tham vấn và giảm hiệu quả hỗ trợ tâm lý.
- Hướng nghiên cứu: Phát triển và cải tiến các thuật toán dựa trên LLMs như LLaMA 3 hoặc Mistral để duy trì tính nhất quán trong phản hồi. Những mô hình này có khả năng lưu trữ và xử lý thông tin từ các cuộc trò chuyện trước, giúp cải thiện sự liên tục và mạch lạc trong đối thoại.
- Bảo mật và quyền riêng tư:
 - Vấn đề: Bảo mật và quyền riêng tư của thông tin người dùng là một mối lo ngại lớn khi sử dụng chatbot cho tham vấn tâm lý. Các dữ liệu nhạy cảm có thể bị lộ hoặc sử dụng sai mục đích nếu không được bảo vệ kỹ lưỡng.
 - Hướng nghiên cứu: Tích hợp các biện pháp bảo mật nâng cao và mã hóa dữ liệu trong các mô hình LLMs như LLaMA 3 hoặc Mistral. Ngoài ra, nghiên cứu các phương pháp phân tích dữ liệu không xâm phạm quyền riêng tư để bảo vệ thông tin cá nhân của người dùng.
- Khả năng xử lý các tình huống khẩn cấp:
 - Vấn đề: Chatbot thường không đủ khả năng để xử lý các tình huống khẩn cấp hoặc nguy hiểm đến tính mạng của người dùng. Điều này có thể dẫn đến hậu quả nghiêm trọng nếu người dùng không nhận được hỗ trợ kịp thời.
 - Hướng nghiên cứu: Sử dụng LLMs như LLaMA 3 hoặc Mistral để phát triển các hệ thống cảnh báo tự động và hướng dẫn người dùng đến các dịch vụ hỗ trợ khẩn cấp. Nghiên cứu cách phát hiện các dấu hiệu của khủng hoảng tâm lý và phản hồi một cách nhanh chóng và hiệu quả.
- Đa dạng văn hóa và ngôn ngữ:
 - Vấn đề: Chatbot hiện tại thường chưa đủ khả năng để đáp ứng nhu cầu của người dùng đến từ các nền văn hóa và ngôn ngữ khác nhau. Điều

này hạn chế phạm vi và hiệu quả của các dịch vụ tham vấn tâm lý trực tuyến.

- Hướng nghiên cứu: Áp dụng các LLMs như LLaMA 3 hoặc Mistral để tạo ra các mô hình ngôn ngữ đa văn hóa và đa ngôn ngữ. Điều này sẽ giúp chatbot hiểu và đáp ứng tốt hơn nhu cầu của người dùng từ nhiều bối cảnh văn hóa và ngôn ngữ khác nhau.

Nhìn chung, việc sử dụng các mô hình ngôn ngữ lớn (LLMs) tiên tiến như LLaMA 3 và Mistral hứa hẹn mang lại những cải tiến đáng kể trong việc giải quyết các vấn đề hiện tại của GenAI chatbot trong tham vấn tâm lý. Các nghiên cứu và phát triển trong tương lai sẽ tập trung vào việc tối ưu hóa các mô hình này để nâng cao hiệu quả và độ tin cậy của các dịch vụ hỗ trợ tâm lý trực tuyến.

1.7 Phương pháp giải quyết

Cùng với sự phát triển nhanh chóng của các mô hình ngôn ngữ lớn, các LLMs ngày càng trở nên hiện đại, có khả năng trả lời các câu hỏi, các tác vụ phức tạp như coding, tư vấn tâm lý, viết content...

Do đó, nhóm quyết định sử dụng các LLMs 7 tỷ tham số phổ biến nhất hiện nay để thực hiện việc fine-tune cho bài toán. Nhóm quyết định sử dụng hai mô hình Vistral và Vinallama cho bài toán, vì đây là hai phiên bản tiếng Việt của hai mô hình Mistral-7b và Llama2-7b. Các mô hình này được pre-trained trên bộ corpus tiếng Việt, cũng như được mở rộng token trong tokenizer để phù hợp hơn với ngôn ngữ tiếng Việt.

Vào ngày 18-4-2024, Meta cho ra mắt mô hình Llama 3 - được cho là một trong những mô hình có performance state-of-the-art, hoạt động tốt ở các bài toán như dịch máy, dialogue generation.... Với tinh thần luôn tìm tòi học hỏi, và cập nhật liên tục và không ngừng thử thách bản thân, nhóm quyết định sẽ fine-tune Llama 3 cho bài toán Tham vấn tâm lý.

CHƯƠNG 2 – CƠ SỞ LÝ THUYẾT

Những nghiên cứu thực nghiệm hoặc lý thuyết: trình bày các cơ sở lý thuyết, lý luận, giả thuyết khoa học và phương pháp nghiên cứu sẽ được sử dụng trong Luận văn, Luận án;

2.1. Mô hình ngôn ngữ lớn

2.1.1 Llama 3

Llama3 là một trong những mô hình ngôn ngữ tiên tiến nhất được phát triển gần đây, đặc biệt tập trung vào khả năng xử lý ngôn ngữ tự nhiên (NLP). Mô hình này được thiết kế để cải thiện khả năng hiểu và tạo ra ngôn ngữ tự nhiên của các hệ thống AI, vượt qua các hạn chế của các phiên bản trước đó như GPT-3 hay BERT.

Llama3 là sản phẩm của sự phát triển không ngừng trong lĩnh vực học sâu và trí tuệ nhân tạo. Được xây dựng trên cơ sở các thành tựu của những mô hình ngôn ngữ trước đó, Llama3 tích hợp nhiều cải tiến quan trọng nhằm tăng cường khả năng hiểu ngữ cảnh và sinh ngữ một cách tự nhiên hơn. Đặc biệt, Llama3 sử dụng kiến trúc Transformer cải tiến và một lượng dữ liệu huấn luyện khổng lồ, giúp mô hình này trở nên mạnh mẽ và linh hoạt hơn.

Các tính năng nổi bật:

- **Khả năng hiểu ngữ cảnh sâu rộng:** Llama3 có khả năng phân tích và hiểu được ngữ cảnh của các đoạn văn bản phức tạp, từ đó đưa ra các phản hồi chính xác và hợp lý hơn.
- **Tạo ra văn bản tự nhiên và mạch lạc:** Với khả năng sinh ngữ được cải thiện, Llama3 có thể tạo ra các đoạn văn bản tự nhiên, mạch lạc và có tính thuyết phục cao, gần giống với văn phong của con người.
- **Khả năng tùy chỉnh và mở rộng:** Llama3 cho phép tùy chỉnh mô hình theo các nhu cầu cụ thể của người dùng, giúp nâng cao hiệu quả ứng dụng trong các lĩnh vực chuyên biệt như y tế, tài chính, giáo dục, và nhiều lĩnh vực khác.
- **Tương tác người dùng tốt hơn:** Nhờ khả năng hiểu biết và tạo ngôn ngữ tự nhiên, Llama3 có thể tương tác với người dùng một cách hiệu quả hơn,

cung cấp các giải pháp hỗ trợ khách hàng, trợ lý ảo, và nhiều ứng dụng khác.

Llama3 có thể được ứng dụng trong nhiều lĩnh vực khác nhau:

- **Chăm sóc khách hàng:** Tạo ra các chatbot thông minh có khả năng giải đáp thắc mắc của khách hàng một cách chính xác và nhanh chóng.
- **Trợ lý ảo:** Hỗ trợ người dùng trong các công việc hàng ngày như lên lịch, gửi email, tìm kiếm thông tin, và nhiều tác vụ khác.
- **Giáo dục:** Hỗ trợ giảng dạy và học tập thông qua việc cung cấp các giải thích chi tiết, tạo ra các bài giảng tự động, và giúp học sinh làm bài tập.
- **Y tế:** Hỗ trợ bác sĩ và bệnh nhân trong việc tra cứu thông tin y khoa, phân tích triệu chứng, và đưa ra các lời khuyên sức khỏe.

2.1.2 Vistral và Vinallama

Vistral (là phiên bản tiếng việt của Mistral-7b, được xây dựng trên nền tảng học sâu và tích hợp các phương pháp xử lý ngôn ngữ tiên tiến nhất. Mô hình này sử dụng một cấu trúc Transformer tùy chỉnh và được huấn luyện trên một tập dữ liệu đa dạng, phong phú, giúp nó có khả năng hiểu và tạo ra văn bản trong nhiều ngữ cảnh khác nhau.

Các tính năng nổi bật của Vistral như sau:

- **Hiệu suất vượt trội:** Vistral được tối ưu hóa để hoạt động hiệu quả hơn, sử dụng ít tài nguyên tính toán hơn mà vẫn duy trì được độ chính xác cao.
- **Độ chính xác cao:** Với khả năng hiểu ngữ cảnh sâu rộng, Vistral có thể tạo ra các phản hồi chính xác và phù hợp hơn trong nhiều tình huống khác nhau.
- **Tính linh hoạt:** Vistral có khả năng điều chỉnh dễ dàng theo các yêu cầu cụ thể của người dùng, từ đó nâng cao hiệu quả trong các ứng dụng thực tế.
- **Tính bảo mật và an toàn:** Vistral được thiết kế với các cơ chế bảo mật tiên tiến, đảm bảo dữ liệu người dùng được bảo vệ và sử dụng một cách an toàn.

Vinallama là phiên bản tiếng Việt của Llama-2-7b, là một mô hình ngôn ngữ tiên tiến được phát triển nhằm cạnh tranh trực tiếp với các mô hình như GPT-4 và Llama3. Với những cải tiến đột phá trong việc xử lý ngôn ngữ tự nhiên, Vinallama

đang trở thành một công cụ mạnh mẽ trong việc tạo ra và hiểu văn bản ngôn ngữ tự nhiên.

Vinallama được phát triển với mục tiêu nâng cao khả năng xử lý ngôn ngữ tự nhiên, tập trung vào việc hiểu ngữ cảnh và tạo ra văn bản một cách tự nhiên và chính xác. Mô hình này sử dụng một phiên bản cải tiến của kiến trúc Transformer và được huấn luyện trên một khối lượng dữ liệu lớn và đa dạng.

Các tính năng nổi bật của Vinallama

- **Khả năng hiểu ngữ cảnh vượt trội:** Vinallama có khả năng phân tích và hiểu được ngữ cảnh phức tạp, từ đó đưa ra các phản hồi chính xác và phù hợp.
- **Tạo ra văn bản tự nhiên và mạch lạc:** Với khả năng sinh ngữ cao, Vinallama có thể tạo ra các đoạn văn bản có tính thuyết phục và tự nhiên, giống như được viết bởi con người.
- **Tính mở rộng và tùy chỉnh:** Vinallama cho phép người dùng tùy chỉnh mô hình theo các nhu cầu cụ thể.

2.2 QLoRA

QLoRA (Quantized Low-Rank Adaptation) là một kỹ thuật tiên tiến trong việc tinh chỉnh các mô hình ngôn ngữ lớn (LLM). Kỹ thuật này nhằm mục đích giảm thiểu chi phí tính toán và tài nguyên cần thiết khi huấn luyện các mô hình lớn, trong khi vẫn duy trì hiệu suất cao.

2.2.1 Khó khăn với mô hình 7 tỷ tham số

Mô hình ngôn ngữ lớn với 7 tỷ tham số mang lại nhiều lợi ích về khả năng xử lý ngôn ngữ tự nhiên, nhưng cũng đi kèm với một số thách thức lớn:

- **Tài nguyên tính toán khổng lồ:** Để huấn luyện một mô hình với 7 tỷ tham số đòi hỏi lượng tài nguyên tính toán rất lớn, bao gồm bộ nhớ GPU và khả năng xử lý cao. Điều này dẫn đến chi phí cao và cần các cơ sở hạ tầng mạnh mẽ.
- **Thời gian huấn luyện dài:** Quá trình huấn luyện mô hình với số lượng tham số lớn có thể mất rất nhiều thời gian, từ vài tuần đến vài tháng, tùy thuộc vào cấu hình phần cứng và kích thước dữ liệu huấn luyện.

- **Khả năng mở rộng và bảo trì:** Quản lý và cập nhật một mô hình lớn như vậy đòi hỏi kỹ thuật phức tạp và kỹ năng cao từ các nhà nghiên cứu và kỹ sư. Việc triển khai và bảo trì mô hình trong môi trường sản xuất cũng là một thách thức lớn.
- **Tiêu thụ năng lượng:** Các mô hình lớn tiêu thụ một lượng năng lượng đáng kể, gây ảnh hưởng đến chi phí vận hành và có thể không bền vững về mặt môi trường.

2.2.2 *QLoRA is all you need*

QLoRA xuất hiện như một giải pháp để khắc phục những khó khăn này. Dưới đây là những điểm chính làm nên tầm quan trọng của QLoRA:

- **Tối ưu hóa tài nguyên:** QLoRA sử dụng kỹ thuật lượng tử hóa và giảm bậc (low-rank adaptation) để giảm số lượng tham số cần thiết trong quá trình huấn luyện. Điều này giúp tiết kiệm tài nguyên tính toán và giảm chi phí.
- **Giữ nguyên hiệu suất cao:** Mặc dù sử dụng ít tham số hơn, QLoRA vẫn duy trì được hiệu suất cao của mô hình gốc. Kỹ thuật này đảm bảo rằng mô hình vẫn có khả năng hiểu và tạo ngôn ngữ tự nhiên một cách chính xác và tự nhiên.
- **Tăng tốc độ huấn luyện:** Bằng cách giảm số lượng tham số và sử dụng các kỹ thuật tối ưu hóa tiên tiến, QLoRA có thể rút ngắn thời gian huấn luyện so với các phương pháp truyền thống. Điều này giúp các nhà nghiên cứu và doanh nghiệp triển khai mô hình nhanh chóng hơn.
- **Khả năng mở rộng:** Với QLoRA, các mô hình có thể dễ dàng mở rộng và điều chỉnh theo các nhu cầu cụ thể mà không cần phải tái huấn luyện toàn bộ mô hình. Điều này mang lại sự linh hoạt và tiện lợi trong việc phát triển và duy trì các ứng dụng AI.

CHƯƠNG 3 – PHƯƠNG PHÁP ĐỀ XUẤT

3.1 Pre-trained các mô hình ngôn ngữ lớn

3.1.1. Lựa chọn mô hình ngôn ngữ lớn

Trong nghiên cứu này, chúng tôi lựa chọn **Llama 3** là mô hình ngôn ngữ lớn (LLM) để tinh chỉnh cho nhiệm vụ chatbot tư vấn tâm lý. Lý do lựa chọn Llama 3 là do:

- **Hiệu năng cao:** Llama 3 được đánh giá là một trong những LLM mạnh mẽ nhất hiện nay, với khả năng xử lý ngôn ngữ tự nhiên tốt, khả năng tạo văn bản chất lượng cao và khả năng tổng hợp thông tin hiệu quả.
- **Khả năng mở rộng:** Llama 3 là một mô hình nguồn mở, cho phép nghiên cứu và tinh chỉnh thêm để phù hợp với các nhiệm vụ cụ thể, trong đó có nhiệm vụ chatbot tư vấn tâm lý.

Tuy nhiên, điểm hạn chế của Llama 3 là mô hình này được đào tạo chủ yếu trên dữ liệu tiếng Anh. Do đó, để đảm bảo Llama 3 có thể hoạt động hiệu quả trong tiếng Việt, chúng tôi sẽ tiến hành bước **pre-training** cho mô hình bằng dữ liệu tiếng Việt.

3.1.2. Mục tiêu của việc pre-training

Mục tiêu của việc pre-training Llama 3 bằng dữ liệu tiếng Việt là:

- **Cải thiện khả năng xử lý ngôn ngữ tiếng Việt:** Giúp Llama 3 hiểu được cấu trúc ngữ pháp, từ vựng, và kiến thức của tiếng Việt, từ đó có thể giao tiếp tự nhiên và chính xác hơn với người dùng tiếng Việt.
- **Nâng cao khả năng hiểu ngữ cảnh:** Bằng cách tiếp xúc với dữ liệu tiếng Việt đa dạng, Llama 3 sẽ học được ngữ cảnh của ngôn ngữ, từ đó hiểu rõ hơn ý nghĩa của các câu hỏi và phản hồi của người dùng.
- **Tăng cường khả năng tạo văn bản tiếng Việt chất lượng:** Việc pre-training giúp Llama 3 học được cách sử dụng các từ ngữ và cấu trúc ngữ pháp tiếng Việt một cách phù hợp, tạo ra các câu trả lời tự nhiên và dễ hiểu cho người dùng.

3.1.3. Kỹ thuật pre-training

Để pre-training Llama 3 bằng dữ liệu tiếng Việt, chúng tôi sẽ sử dụng kỹ thuật **qLoRA**. qLoRA là một kỹ thuật tinh chỉnh LLM hiệu quả, cho phép chúng ta tinh chỉnh LLM với một lượng bộ nhớ hạn chế. Kỹ thuật này giúp giảm thiểu thời gian và chi phí tính toán cần thiết để pre-training.

Kết luận, việc lựa chọn Llama 3 làm LLM chính và sử dụng kỹ thuật pre-training bằng qLoRA là một cách tiếp cận hiệu quả để tạo ra một chatbot tư vấn tâm lý hiệu quả cho người dùng tiếng Việt. Bằng cách tiếp cận này, chúng tôi mong muốn chatbot sẽ hiểu được ngôn ngữ tiếng Việt một cách chính xác và tự nhiên hơn, từ đó cung cấp dịch vụ tư vấn tâm lý chất lượng cho người dùng.

3.2 Instruction Fine-tune cho bài toán chatbot

3.2.1. Lý do lựa chọn kỹ thuật hướng dẫn

Sau khi pre-training Llama 3 bằng dữ liệu tiếng Việt, chúng ta cần tinh chỉnh thêm mô hình để nó có thể thực hiện tốt nhiệm vụ chatbot tư vấn tâm lý. Kỹ thuật hướng dẫn (Instruction Fine-tuning) là lựa chọn phù hợp vì lý do sau:

- **Cải thiện khả năng hiểu và thực thi hướng dẫn:** Kỹ thuật hướng dẫn giúp LLM hiểu rõ hơn ý định của người dùng và thực thi các yêu cầu một cách chính xác.
- **Tăng cường khả năng giải quyết các vấn đề phức tạp:** Kỹ thuật hướng dẫn cho phép LLM giải quyết các vấn đề phức tạp hơn, bao gồm cả những nhiệm vụ đòi hỏi sự logic, sáng tạo và khả năng suy luận.
- **Nâng cao tính linh hoạt của chatbot:** Kỹ thuật hướng dẫn giúp chatbot trở nên linh hoạt hơn, có thể xử lý nhiều loại yêu cầu khác nhau từ người dùng.

3.2.2. Kỹ thuật instruction fine-tuning

Để thực hiện instruction fine-tuning, chúng tôi sẽ sử dụng kỹ thuật **qLoRA**. Kỹ thuật này được lựa chọn vì khả năng tinh chỉnh mô hình một cách hiệu quả với lượng bộ nhớ hạn chế.

3.2.3. *Áp dụng kỹ thuật instruction fine-tuning cho các mô hình*

Chúng tôi sẽ áp dụng kỹ thuật instruction fine-tuning cho 3 mô hình:

- **Llama 3:** Mô hình Llama 3 đã được pre-training bằng dữ liệu tiếng Việt như đã trình bày ở phần 3.1. Việc instruction fine-tuning giúp Llama 3 học cách thực hiện các nhiệm vụ chatbot tư vấn tâm lý một cách hiệu quả.
- **Vistral:** Mô hình Vistral là một LLM được đào tạo bởi VietAI, đã được tối ưu hóa cho tiếng Việt. Việc instruction fine-tuning giúp Vistral nâng cao khả năng xử lý các yêu cầu cụ thể của chatbot tư vấn tâm lý.
- **Vinallama:** Mô hình Vinallama cũng là một LLM được đào tạo bởi VietAI, chuyên biệt cho tiếng Việt. Việc instruction fine-tuning giúp Vinallama học cách tạo ra các phản hồi phù hợp với ngữ cảnh của chatbot tư vấn tâm lý.

Việc lựa chọn các mô hình Vistral và Vinallama là do chúng đã được pre-trained với dữ liệu tiếng Việt và có khả năng xử lý ngôn ngữ tiếng Việt tốt hơn so với Llama 3 chưa được pre-trained. Instruction fine-tuning sẽ giúp các mô hình này đạt hiệu quả cao hơn trong nhiệm vụ chatbot tư vấn tâm lý.

3.3 Lượng tử hóa (quantization) mô hình ngôn ngữ lớn

3.3.1. *Cần lượng tử hóa mô hình*

Các mô hình ngôn ngữ lớn (LLM) như Llama 3 (8 tỷ tham số), Vistral (7 tỷ tham số), Vinallama (7 tỷ tham số) có kích thước rất lớn, đòi hỏi lượng bộ nhớ và sức mạnh tính toán cao để chạy. Do đó, việc chạy các mô hình này trên máy tính cá nhân (local) với GPU thông thường như Google Colab T4 GPU là không khả thi.

3.3.2. *Lượng tử hóa là gì và mục đích*

Lượng tử hóa là một kỹ thuật nén mô hình máy học, bao gồm việc giảm kích thước của các tham số trong mô hình từ dạng số thực (float) sang dạng số nguyên (integer).

Mục tiêu của lượng tử hóa là:

- **Giảm kích thước mô hình:** Giúp giảm lượng bộ nhớ cần thiết để lưu trữ mô hình, từ đó cho phép chạy mô hình trên các thiết bị có bộ nhớ hạn chế hơn.
- **Tăng tốc độ tính toán:** Việc sử dụng số nguyên thay cho số thực giúp giảm thiểu lượng tính toán, dẫn đến tăng tốc độ xử lý.

3.3.3. Lợi ích của lượng tử hóa

Việc lượng tử hóa mang lại một số lợi ích quan trọng:

- **Giảm kích thước mô hình:** Giúp giảm dung lượng lưu trữ, cho phép triển khai mô hình trên các thiết bị có bộ nhớ hạn chế.
- **Tăng tốc độ tính toán:** Giúp chạy mô hình nhanh hơn, đặc biệt trên các thiết bị có khả năng tính toán thấp.
- **Giảm chi phí triển khai:** Giảm chi phí tính toán và lưu trữ, giúp cho việc triển khai mô hình trở nên dễ dàng hơn.

3.3.4. Áp dụng lượng tử hóa cho các mô hình

Để giải quyết vấn đề về kích thước mô hình lớn, chúng tôi quyết định áp dụng kỹ thuật lượng tử hóa cho các mô hình Llama 3, Vistral và Vinallama.

Sau khi lượng tử hóa, các mô hình được chuyển đổi sang định dạng **gguf**. Định dạng gguf cho phép chúng tôi chạy các mô hình trên CPU, sử dụng các phần mềm như LM Studio, mà không cần GPU mạnh mẽ. Việc này giúp đơn giản hóa quá trình triển khai và thử nghiệm chatbot.

Việc lựa chọn các mô hình Llama 3, Vistral và Vinallama cùng với việc áp dụng kỹ thuật instruction fine-tuning và lượng tử hóa là một phương pháp tiếp cận hiệu quả để tạo ra một chatbot tư vấn tâm lý hoạt động hiệu quả trên thiết bị local. Bằng cách này, chúng tôi có thể triển khai và thử nghiệm chatbot một cách linh hoạt và tiết kiệm chi phí.

3.4 Dữ liệu

3.4.1. Nguồn dữ liệu

Bộ dữ liệu được sử dụng để huấn luyện chatbot tư vấn tâm lý được thu thập từ các nguồn sau:

- **Diễn đàn sinh viên:** Đây là diễn đàn dành riêng cho sinh viên Tôn Đức Thắng, tại đây sẽ có những câu hỏi, những tâm sự của sinh viên Tôn Đức Thắng
- **Diễn đàn tâm lý:** Các diễn đàn tâm lý trực tuyến cung cấp một lượng lớn dữ liệu về các cuộc thảo luận và tư vấn tâm lý, bao gồm cả câu hỏi, phản hồi và chia sẻ kinh nghiệm của người dùng.
- **Hugging Face:** Nền tảng Hugging Face là kho lưu trữ dữ liệu và mô hình máy học lớn, cung cấp nhiều bộ dữ liệu liên quan đến xử lý ngôn ngữ tự nhiên, trong đó có các bộ dữ liệu về tâm lý học.
- **Kaggle:** Nền tảng Kaggle là một cộng đồng dành cho các nhà khoa học dữ liệu, cung cấp các bộ dữ liệu mở liên quan đến nhiều lĩnh vực, bao gồm cả dữ liệu về tâm lý học.

3.4.2. Tiền xử lý dữ liệu

Dữ liệu ban đầu được thu thập chủ yếu bằng tiếng Anh. Để phù hợp với mục tiêu của nghiên cứu, nhóm nghiên cứu quyết định dịch dữ liệu sang tiếng Việt bằng các công cụ dịch máy như Gemini, GPT 3.5, Google Translate, v.v. Việc sử dụng nhiều công cụ dịch máy nhằm tạo ra một bộ dữ liệu với phong cách ngôn ngữ đa dạng.

3.4.3. Các loại dữ liệu

Ngoài dữ liệu về tham vấn tâm lý, nhóm nghiên cứu sử dụng thêm hai loại dữ liệu:

- **Dữ liệu Alpaca:** Bộ dữ liệu Alpaca là một tập hợp các dữ liệu hướng dẫn (instruction) và phản hồi tương ứng, được sử dụng để huấn luyện LLM có khả năng tuân theo hướng dẫn và tạo ra các phản hồi phù hợp với yêu cầu.

- **Mục đích sử dụng:** Dữ liệu Alpaca được sử dụng để giúp chatbot học cách thực hiện các nhiệm vụ cụ thể theo yêu cầu của người dùng. Ví dụ: chatbot có thể học cách giải thích một khái niệm, tóm tắt một bài viết, hoặc viết một bài thơ.
- **Dữ liệu English General Knowledge:** Bộ dữ liệu này cung cấp kiến thức chung về tiếng Anh.
 - **Mục đích sử dụng:** Dữ liệu English General Knowledge được sử dụng để ngăn chặn hiện tượng **language drift** - tình trạng LLM bị "lệch" ngôn ngữ và tạo ra các phản hồi không phù hợp với ngữ cảnh.

CHƯƠNG 4: KẾT QUẢ THỰC NGHIỆM

4.1 Dữ liệu

4.1.1 Dữ liệu cho text completion fine-tuning

Đối với bộ dữ liệu corpus tiếng Việt, nhóm đã sử dụng hai bộ dữ liệu đó là: **Binhvq news**: Đây là corpus tiếng Việt nổi tiếng, được sử dụng để pre-trained các mô hình tiếng Việt phổ biến như Vinallam, phoGPT,...

- Trích xuất từ khoảng 14.896.998 bài báo trên internet bao gồm các báo
- Link dữ liệu: <https://github.com/binhvq/news-corpus>
- bao gồm: 19.4M dòng
- kích thước: 4.78GB
- **Vi_wikipedia**: Dữ liệu được crawl về từ các trang wikipedia tiếng Việt
- Link dữ liệu: https://huggingface.co/datasets/vietgpt/wikipedia_vi
- Bao gồm: 1M dòng
- kích thước: 570M

4.1.2 Dữ liệu tham vấn tâm lý

Đối với dữ liệu tham vấn tâm lý, nhóm đã sử dụng bộ dữ liệu về tham vấn tâm lý trên Huggingface. Tuy nhiên, có một vấn đề đó là bộ dữ liệu trên là tiếng Anh.

Do đó, nhóm quyết định sẽ dịch 80% dữ liệu ra thành tiếng Việt, với sự hỗ trợ của Google Translate API, Gemini API, Groq API, và ChatGPT. Việc sử dụng nhiều công cụ để dịch giúp cho dữ liệu tiếng Việt có được văn phong đa dạng, không quá rập khuôn theo một khuôn khổ của từng công cụ.

Tóm lại, bộ dữ liệu về tham vấn gồm có:

- 80% tiếng Việt, được dịch từ các công cụ trên
- 20% tiếng Anh, bộ dữ liệu gốc

Ngoài ra, bên cạnh dữ liệu là các dataset về tham vấn tâm lý. Nhóm tiếp tục trộn bộ dữ liệu với hai bộ dữ liệu khác:

Vi_alpaca: bộ dữ liệu alpaca là một bộ dataset nổi tiếng để thực hiện instruction fine-tune cho các mô hình llama. Bộ dữ liệu này bao gồm các instruction,

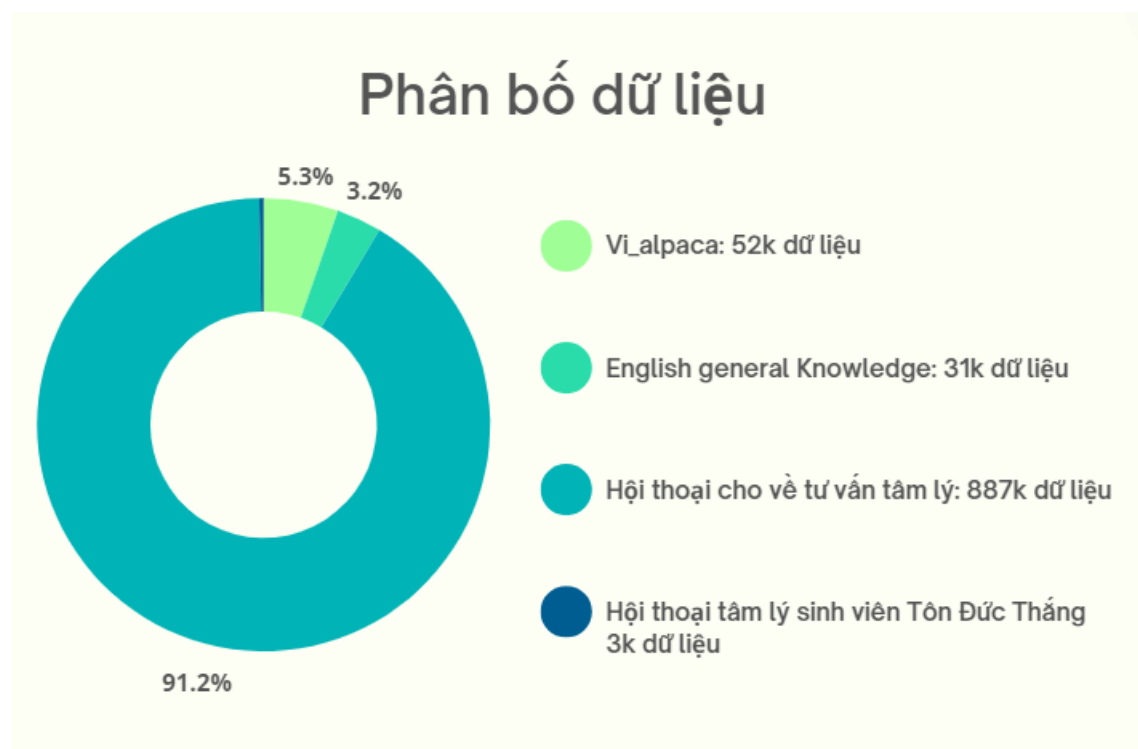
input, output về đủ các chủ đề, giúp mô hình học được cách thực hiện các tác vụ khác nhau, như viết code đơn giản, tóm tắt văn bản,... Vi_alpaca là phiên bản tiếng Việt của Alpaca, được dịch bằng GPT4.

Bộ dữ liệu có 52 nghìn dòng

English general knowledge: Là bộ dữ liệu hỏi đáp về các kiến thức phổ thông bằng tiếng Anh.

Bộ dữ liệu bao gồm 31 nghìn dòng

Mục tiêu của bộ dữ liệu này là: giúp mô hình tránh được hiện tượng Language drift



4.2 Fine-tune

4.2.1 Text completion fine-tuning

Công đoạn Completion fine-tuning được thực hiện với mô hình llama 3, Mục tiêu giúp mô hình học được cấu trúc ngữ pháp, từ vựng, cũng như là các kiến thức tiếng Việt.

Ở giai đoạn này, nhóm thực hiện việc fine-tune với kỹ thuật QLoRA, với các Hyperparameter sau:

- maximum input length: 2000
- rank: 8
- lora alpha: 8
- batch_size: 8
- num_epoch: 1
- GPU: T4
- Số lượng tham số:

Quá trình fine-tune đạt được những kết quả sau:

- thời gian: 2 tiếng
- training_loss: 2.6606

4.2.2 Instruction fine-tuning

Ở giai đoạn này, nhóm đã fine-tune trên cả ba mô hình: Llama 3 (sau khi fine-tune với corpus tiếng Việt), Vistral, và Vinallama

Prompt template sử dụng để fine-tune cho mô hình:

Nếu bạn là một chuyên gia tâm lý với nhiều năm kinh nghiệm, nhiệm vụ của bạn là cung cấp những lời khuyên, giải pháp, nhằm nâng cao chất lượng cuộc sống của mọi người.

`</start_header_id/>user</end_header_id/>`

`{}`

`</eot_id/>`

`</start_header_id/>user</end_header_id/>`

`{}`

`</eot_id/>`

Đối với các tham số fine-tune cũng tương tự như trên, ở cả ba mô hình, nhóm đều sử dụng QLoRA:

- rank: 8
- batch_size: 8
- num_epoch: 1
- max input length: 2000

Các kết quả đạt được đối với llama 3:

- thời gian fine-tune:
- training_loss

Vistral:

- thời gian:
- training_loss:

Vinallama:

- thời gian:
- training_loss:

4.2.3 Kết quả đạt được

Để đánh giá mô hình, nhóm tạo bộ dữ liệu gồm 50 câu hỏi và câu trả lời tương ứng với mỗi mô hình. Các câu hỏi thuộc 5 nhóm trong lĩnh vực tham vấn tâm lý học đường:

- General: các câu hỏi dạng chung chung như ‘Trầm cảm là gì?’, ‘dấu hiệu của trầm cảm’,...
- Personal loss: các vấn đề liên quan đến mất mát, như sự ra đi của người thân, hay vật nuôi, hay mất đi một người bạn thân
- Relationship: các vấn đề về các mối quan hệ cá nhân, như người yêu, bạn bè, gia đình....
- Academic: các vấn đề liên quan đến học tập như áp lực học hành, thi cử...

- Career: các nỗi lo của sinh viên về vấn đề việc làm

Các mô hình sẽ thực hiện trả lời các câu hỏi trên, các tham số cho việc inference:

- max_new_token: 2000
- temperature: 0.8,
- top_k: 20
- repetition_penalty: 1.2

Việc đánh giá được thực hiện bằng cách sử dụng GPT-4o - mô hình có performance tốt nhất hiện nay, thực hiện việc đánh giá các cặp câu hỏi và trả lời, dựa trên 4 tiêu chí:

- Helpfulness: How useful is the answer?
- Relevance: How relevant is the information?
- Accuracy: How accurate is the information?
- Level of Detail: How detailed is the answer?

Việc đánh giá được thực hiện với prompt sau:

Please rate the helpfulness, relevance, accuracy, and level of detail of the AI assistants' responses. Each assistant receives a score for each criterion on a scale of 1 to 10, where a higher score indicates better performance.

Evaluation Criteria

Helpfulness: How useful is the answer?

Relevance: How relevant is the information?

Accuracy: How accurate is the information?

Level of Detail: How detailed is the answer?

Example

Question: Tôi cảm thấy mệt mỏi vì mâu thuẫn liên tục với bạn trai/bạn gái, tôi nên làm gì?

Assistant's Answer: Có vẻ như bạn đang trải qua một giai đoạn khó khăn trong mối quan hệ của mình. Có điều gì cụ thể mà bạn muốn thay đổi về tình hình hiện tại không?

```

Scores:

Helpfulness: 8
Relevance: 9
Accuracy: 8
Level of detail: 7
User Question and Assistant's Answer

---
{}
---

Scores:
Helpfulness:
Relevance:
Accuracy:
Level of detail:
Please provide your scores based on the evaluation criteria above.
Only return scores, no extra explanation and text, except the order
of question-answer pair

```

Kết quả đạt được:

Question type	samples	Vinallama	Vistra l	Llama3
General	10	77.5	85.75	81.25
Relationship	10	79.25	77.5	78.5
Personal loss	10	71.25	74.75	84
Academic pressure	10	77.25	79	82
Career	10	67.5	73.75	78
Total	50	74.55	78.15	80.625

Từ bảng trên, ta thấy được, llama 3 đạt hiệu suất cao hơn so với 2 mô hình còn lại, trong khi Vistral thì outperform ở các câu hỏi general, thì Vinallama lại trả lời tốt các câu hỏi về relationship.

4.3 Xây dựng chatbot

4.3.1 Ollama

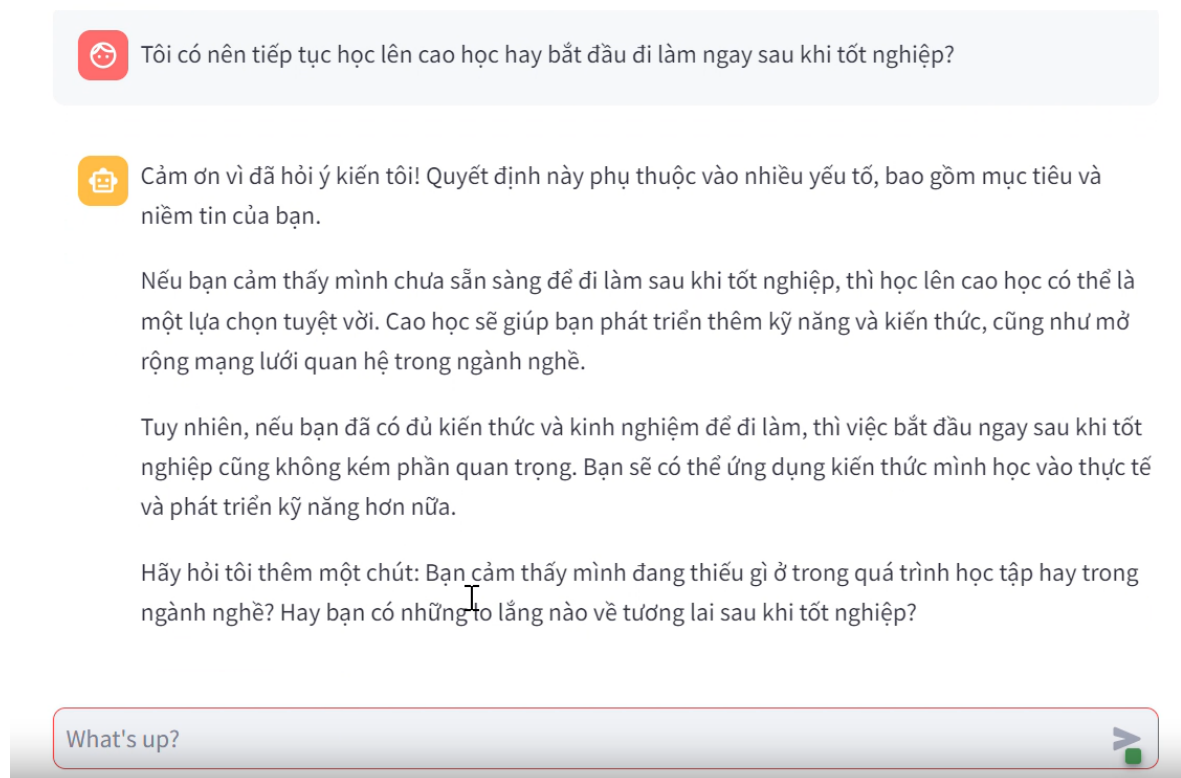
Để có thể config mô hình với Ollama, các model sau khi fine-tune sẽ được lượng tử hóa (quantization) xuống còn 4bit mỗi parameter và chuyển thành file gguf, giúp cho mô hình nhẹ hơn, nhưng vẫn giữ được độ chính xác khá cao.

Ollama là một công cụ được thiết kế để chạy các mô hình AI dưới local. Ollama hỗ trợ đa nền tảng từ MacOS, Windows cho tới Linux. Nó cho phép bạn khai thác sức mạnh của AI mà không phụ thuộc vào các nhà cung cấp bên ngoài như OpenAI, Google...

4.3.2 Streamlit

Streamlit là một thư viện mã nguồn mở, giúp xây dựng các application với Python một cách nhanh chóng. Do hạn chế về mặt thời gian, nên nhóm quyết định chọn streamlit.

Dưới đây là một số ví dụ kết quả trả lời của chatbot:



So sánh với chat GPT trả lời:

ChatGPT ▾


tôi có nên tiếp tục học lên cao học hay bắt đầu đi làm ngay sau khi tốt nghiệp?

Việc quyết định tiếp tục học lên cao học hay bắt đầu đi làm ngay sau khi tốt nghiệp là một quyết định quan trọng và phụ thuộc vào nhiều yếu tố cá nhân. Dưới đây là một số điểm cần xem xét để giúp bạn đưa ra quyết định:

Lợi ích của việc học lên cao học:

- Chuyên môn sâu hơn:** Cao học giúp bạn có kiến thức chuyên sâu và kỹ năng cao hơn trong lĩnh vực của mình.
- Cơ hội việc làm tốt hơn:** Bằng cấp cao hơn có thể giúp bạn cạnh tranh hơn trên thị trường lao động và có nhiều cơ hội việc làm tốt hơn, với mức lương cao hơn.
- Phát triển cá nhân và nghiên cứu:** Nếu bạn yêu thích nghiên cứu và học hỏi, cao học sẽ cung cấp cho bạn cơ hội để thực hiện những nghiên cứu sâu sắc hơn và phát triển bản thân.
- Mạng lưới quan hệ:** Học lên cao học có thể giúp bạn mở rộng mạng lưới quan hệ trong ngành của mình, gặp gỡ và hợp tác với các chuyên gia và học giả hàng đầu.

Chat GPT đánh giá câu trả lời mà chatbot trả lời như trên:

 **Đánh giá chi tiết với điểm cụ thể**

- 1. Helpfulness (Mức độ hữu ích)**
 - Điểm: 9/10
- 2. Relevance (Mức độ liên quan)**
 - Điểm: 10/10
- 3. Accuracy (Độ chính xác)**
 - Điểm: 9/10
- 4. Level of Detail (Mức độ chi tiết)**
 - Điểm: 8/10



Tôi cảm thấy mình thiếu nhiều kinh nghiệm thực tế, những thứ tôi học khác xa với những thứ doanh nghiệp cần, do đó rất khó để xin việc, nên tôi cũng đang cân nhắc việc học thạc sĩ



Thưa bạn! Cảm ơn vì đã chia sẻ với tôi. Thân thiết kinh nghiệm thực tế là một nhu cầu thiết yếu trong thị trường việc làm hiện nay. Học thạc sĩ có thể là một lựa chọn tuyệt vời để giúp bạn bổ sung những thiếu sót về kinh nghiệm và kiến thức. Tuy nhiên, trước khi quyết định học thạc sĩ, bạn nên cân nhắc về mục tiêu của mình và xem liệu có bao nhiêu người trong ngành nghề này có bằng thạc sĩ. Cũng như vậy, bạn nên xem xét lại nhu cầu thực tế của doanh nghiệp và những gì họ cần từ một ứng viên. Bạn có thể hỏi thêm một chút: Bạn đã suy nghĩ về những gì bạn muốn đạt được sau khi học thạc sĩ? Hay bạn có những lo lắng nào về việc học thạc sĩ sẽ giúp bạn cải thiện tình hình công việc?

What's up?

4.4 Ý nghĩa kết quả nghiên cứu

4.4.1 Ý nghĩa khoa học

- Xây dựng bộ dữ liệu về tư vấn tâm lý cho sinh viên, gồm các bộ câu hỏi, tình huống thực tế.
- Phát triển được mô hình học máy, học sâu, mô hình ngôn ngữ lớn phù hợp với bài toán trên, cũng như là hoạt động tốt với bộ dữ liệu tiếng Việt.

4.4.2 Ý nghĩa thực tiễn

- Tư vấn tâm lý cho thế hệ sinh viên trẻ Việt Nam, đưa ra các câu trả lời chất lượng, giúp các bạn trẻ có thể định hướng được trong cuộc sống, truyền cảm hứng, và vượt qua được lo âu.
- Giúp tiết kiệm chi phí tư vấn tâm lý, giảm bớt gánh nặng về tài chính cho sinh viên.
- Phân tích, đánh giá, dự đoán được các vấn đề tâm lý mà sinh viên hay gặp phải, từ đó đưa ra hướng giải quyết hợp lý và kịp thời.

4.4 Đóng góp của nghiên cứu

- Mô hình llama 3 tiếng Việt, có thể tiếp tục fine-tune cho các tác vụ khác với tiếng Việt như chatbot chăm sóc khách hàng, dịch máy, tóm tắt văn bản,...

- Các mô hình được fine-tune cho bài toán chatbot
- Chatbot tham vấn tâm lý

4.5 Những điều đạt được

- Nghiên cứu được các mô hình ngôn ngữ lớn hiện đại cho cả tiếng Anh lẫn tiếng Việt: Llama3, Vistral, Vinallama
- Áp dụng được kĩ thuật fine-tune với QLoRA, giúp thực hiện fine-tune nhanh, tiết kiệm được tài nguyên phần cứng, đạt hiệu suất cao
- Xây dựng được mô hình phù hợp cho tác vụ tham vấn tâm lý
- Xây dựng được giao diện chat đơn giản, dễ sử dụng

4.6 Định hướng nghiên cứu tiếp theo

- Tiếp tục fine-tune với RLHF, DPO - các kĩ thuật reinforcement learning tiên tiến hiện nay để phù hợp hơn với bài toán, bối cảnh, hoàn cảnh, tránh các câu trả lời tiêu cực, gây ảnh hưởng xấu đến người dùng.
- Xây dựng giao diện hoàn chỉnh, có hệ thống đăng ký đăng nhập, cá nhân hóa cho từng người dùng
- Phát triển sản phẩm hoàn chỉnh, sẵn sàng đưa ra thị trường.

TÀI LIỆU THAM KHẢO

Tiếng Anh

1. Introducing Meta Llama 3: The most capable openly available LLM to date
<https://ai.meta.com/blog/meta-llama-3/>
2. Jiang, A. Q., Sablayrolles, A., Mensch, A., Bamford, C., Chaplot, D. S., Casas, D. D. L., ... & Sayed, W. E. (2023). Mistral 7B. *arXiv preprint arXiv:2310.06825*.
3. Nguyen, Q., Pham, H., & Dao, D. (2023). VinaLLaMA: LLaMA-based Vietnamese Foundation Model. *arXiv preprint arXiv:2312.11011*.
4. Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., ... & Wen, J. R. (2023). A survey of large language models. *arXiv preprint arXiv:2303.18223*.
5. Doan, V. T., Truong, Q. T., Nguyen, D. V., Nguyen, V. T., & Luu, T. N. N. (2023, October). Efficient Finetuning Large Language Models For Vietnamese Chatbot. In *2023 International Conference on Multimedia Analysis and Pattern Recognition (MAPR)* (pp. 1-6). IEEE.