ỨNG DỤNG MÔ HÌNH HỌC MÁY VÀ XỬ LÝ NGÔN NGỮ TỰ NHIÊN TRONG VIỆC XÂY DỰNG CHATBOT HỖ TRỢ XÁC ĐỊNH VÀ CHỮA LÀNH TÂM LÝ HẬU THƯƠNG TỔN

Đoàn Duy Thành*, Nguyễn Ngọc Duy
Phạm Hoàng Đức Tài, Đỗ Thị Thanh Thảo
Nguyễn Thị Nhung, Phạm Thị Kim Phượng, Nguyễn Thái Cường
Trường Đại học Công nghiệp Hà Nội

Tóm tắt

Theo kết quả khảo sát của nhóm tác giả, phần đông sinh viên đã và đang gặp phải các vấn đề chấn thương tâm lý có thể dẫn đến nhiều vấn đề sức khỏe tâm thần nghiêm trọng, ảnh hưởng tiêu cực đến chất lượng cuộc sống có thể dẫn đến trầm cảm, rối loạn lưỡng cực,... nếu kéo dài và không được chữa trị có thể khiến người bệnh có những quyết định thiếu sáng suốt như gây tổn thương chính bản thân hoặc người khác. Nghiên cứu này đề xuất ứng dụng mô hình học máy và xử lý ngôn ngữ tự nhiên (NLP) để xây dựng chatbot hỗ trợ xác định và chữa lành tâm lý hậu thương tổn. Chatbot sẽ sử dụng các kỹ thuật như: Phân loại văn bản, học máy, xử lý ngôn ngữ tự nhiên. Ứng dụng được cấu hình để phù hợp cho đối tượng mắc phải các hội chứng tâm lý.

Từ khóa: Chấn thương tâm lý; Xử lý ngôn ngữ tự nhiên; Học máy; Chatbot.

Abstract

Application of machine learning models and natural language processing in building chatbots to support the identification and healing of post-traumatic psychological wounds.

Nowadays, students in general have to cope with psychological trauma issues caused by the impact of rare events in life. Psychological trauma can lead to many serious mental health issues, negatively affecting the quality of life, and potentially leading to depression and bipolar disorder,... if prolonged, it can lead patients to make irrational decisions, causing harm to themselves or others. This study proposes the application of machine learning models and Natural Language Processing (NLP) to build a chatbot to support the identification and healing of post-traumatic psychological trauma. The chatbot will use techniques such as text classification, machine learning, and natural language processing. The application is configured to be suitable for individuals with post-traumatic psychological trauma.

Keywords: Psychologica; Natural language processing; Machine learning; Chatbot.

***Tác giả liên hệ, Email:** thanhdd@st.haui.edu.vn

DOI: https://doi.org/10.63064/khtnmt.2024.566

1. Giới thiệu

Tâm lý học là một chủ đề cực kỳ quan trọng và nhạy cảm trong cuộc sống con người, việc tổn thương tâm lý cũng là một dạng thức khác của việc bị bệnh thường gặp như cảm, sốt, gãy tay, gãy chân hay ung thư,… tất cả đều khiến cho người mắc phải gặp những bất lợi trong cuộc sống và sinh hoạt. Ngày nay, tỷ lệ trầm cảm, rối loạn lưỡng cực hay các dạng hội chứng tâm lý khác đang trở nên cực kỳ phổ biến ở lứa tuổi sinh viên. Theo khảo sát của nhóm nghiên cứu được tiến hành trong tháng 10 năm 2023 với 385 sinh viên tại thư viện Trường Đại học Công nghiệp Hà Nội, khảo sát có độ tin cậy đạt 95 % cho thấy 85 % số sinh viên khảo sát cho rằng bản thân gặp phải các sự kiện dẫn đến thương tổn, 89 % trong số đó mắc một hoặc đa hội chứng tâm thần khác nhau. Xã hội phát triển cùng với sự phát triển của mạng xã hội, khoảng cách thế hệ Gen Z với các thế hệ khác phần nào đó khiến cho sự tiếp xúc giữa người với người một cách đủ sâu sắc dần trở nên khó khăn hơn, bằng chứng là trong khảo sát chưa tới 27 % số sinh viên có người để họ tin tưởng và chia sẻ vấn đề kể cả là bạn bè hay gia đình. Điều này dẫn đến sự cô lập trong tâm lý nếu không được định hướng đúng đắn sẽ dễ dẫn tới những hành vi không đáng có về sau. Ngoài ra theo khảo sát cho thấy 71 % số sinh viên cảm thấy ngại và không thoải mái khi phải chia sẻ về vấn đề của mình, điều này sẽ dần tích tụ lại và cuối cùng bộc phát bằng các hành động hay tổn thương không thể chữa lành.

Ngày nay việc xác nhận các loại tổn thương tâm lý gần như đều nhờ vào các bác sĩ tâm lý học, tuy nhiên hạn chế về mặt kinh phí cũng như ngại chia sẻ vấn đề của mình với người khác khiến cho cách làm này trở nên hạn chế. Các ứng dụng hiện nay đa phần là đều nằm ở dạng tích chọn rồi đưa ra điểm số, chưa giải quyết được vấn đề được kết nối, từ đó không thể có cái nhìn sâu sắc và chính xác về những tổn thương của bệnh nhân.

Với sự phát triển về công nghệ, đặc biệt là trong lĩnh vực xử lý ngôn ngữ tự nhiên và học máy đã giúp cho việc xây dựng các ứng dụng chatbot có sự phản hồi tự nhiên trở nên thực tế hơn rất nhiều. Với xử lý ngôn ngữ tự nhiên (NLP) các đoạn văn bản sẽ được hỗ trợ xác định tìm hiểu ý nghĩa của câu hỏi, phản hồi từ người dùng một cách chính xác hơn, chatbot sẽ dễ dàng hiểu và đánh giá được xem người dùng đang có những trạng thái như nào để có thể đưa ra được hướng tiếp cận phù hợp. Các mô hình học máy cũng đóng vai trò quan trọng trong việc nhận biết và phản hồi tự động các tình huống tâm lý khác nhau.

Điểm mới và thực tiễn của nghiên cứu là ứng dụng mô hình học máy và xử lý ngôn ngữ tự nhiên trong việc xây dựng chatbot hỗ trợ xác định và chữa lành tâm lý hậu thương tổn này nằm ở chỗ dần đưa công nghệ tiếp cận đến cảm xúc và trạng thái của con người, giúp giải quyết một bài toán cực kỳ nhạy cảm liên quan đến xác định các hội chứng tâm lý. Đồng thời nghiên cứu này cũng có tiềm năng để phát triển trong tương lai khi chúng ta sẽ dần có cái nhìn đúng đắn hơn về bệnh tâm lý.

2. Cơ sở lý thuyết

2.1. Xử lý ngôn ngữ tự nhiên và trích chọn thuộc tính với TF-IDF [3]:

Xử lý ngôn ngữ tự nhiên (NLP) là một nhánh của trí tuệ nhân tạo tập trung vào các ứng dụng trên ngôn ngữ của con người. Nhóm tác giả phân tích các văn bản ngôn ngữ Anh.

Các bước xử lý của NLP bao gồm:

- Phân tích hình thái - trong bước này từng từ sẽ được phân tích và các ký tự không phải chữ (như các dấu câu) sẽ được tách ra khỏi các từ. Trong tiếng Anh và nhiều ngôn ngữ khác, các từ được phân tách với nhau bằng dấu cách. Tuy nhiên trong tiếng Việt, dấu cách được dùng để phân tách các tiếng (âm tiết) chứ không phải từ. Cùng với các ngôn ngữ như tiếng Trung, tiếng Hàn, tiếng Nhật, phân tách từ trong tiếng Việt là một công việc không hề đơn giản.

- Phân tích cú pháp - dãy các từ sẽ được biến đổi thành các cấu trúc thể hiện sự liên kết giữa các từ này. Sẽ có những dãy từ bị loại do vi phạm các Luật văn phạm.

- Phân tích ngữ nghĩa - thêm ngữ nghĩa vào các cấu trúc được tạo ra bởi bộ phân tích cú pháp.

- Tích hợp văn bản - ngữ nghĩa của một câu riêng biệt có thể phụ thuộc vào những câu đứng trước, đồng thời nó cũng có thể ảnh hưởng đến các câu phía sau.

- Phân tích thực nghĩa - cấu trúc thể hiện điều được phát ngôn sẽ được thông dịch lại để xác định nó thật sự có nghĩa là gì.

TF-IDF: Trong khai phá dữ liệu văn bản (Text mining), thuật ngữ TF-IDF (Term Frequency - Inverse Document Frequency) là một phương thức thống kê được biết đến rộng rãi nhất để xác định độ quan trọng của một từ trong đoạn văn bản trong một tập nhiều đoạn văn bản khác nhau. Nó thường được sử dụng như một trọng số trong việc khai phá dữ liệu văn bản. TF-IDF chuyển đổi dạng biểu diễn văn bản thành dạng không gian vector (VSM), hoặc thành những vector thưa thớt.

- TF (Term Frequency): Là tần suất xuất hiện của một từ trong một đoạn văn bản. Với những đoạn văn bản có độ dài khác nhau, sẽ có những từ xuất hiện nhiều ở những đoạn văn bản dài thay vì những đoạn văn bản ngắn. Vì thế, tần suất này thường được chia cho độ dài của đoạn văn bản như một phương thức chuẩn hóa (Normalization). TF được tính bởi công thức:

$$tf(t,d) = \frac{f\,(t,d)}{\max\{f(w,d) : w \in d\}}$$

(với t là một từ trong đoạn văn bản; f (t, d) là số lần xuất hiện của t trong đoạn văn bản d, tf (t, d) là tần suất xuất hiện của từ t trong văn bản d và max {f(w, d):w∈d} số lần xuất hiện của từ có số lần xuất hiện nhiều nhất trong văn bản d).

- IDF (Inverse Document Frequency): Tính toán độ quan trọng của một từ. Khi tính toán TF, mỗi từ đều quan trọng như nhau, nhưng có một số từ trong tiếng Anh như *"is", "of", "that",...* xuất hiện khá nhiều nhưng lại rất ít quan trọng. Vì vậy, chúng ta cần một phương thức bù trừ những từ xuất hiện nhiều lần và tăng độ quan trọng của những từ ít xuất hiện nhưng có ý nghĩa đặc biệt cho một số đoạn văn bản hơn bằng cách tính IDF:

$$idf(t) = \frac{N}{t \in D : t \in d}$$

(trong đó *N* là tổng số đoạn văn bản; tập |{t \in D : t \in d}| là số văn bản chứa từ).

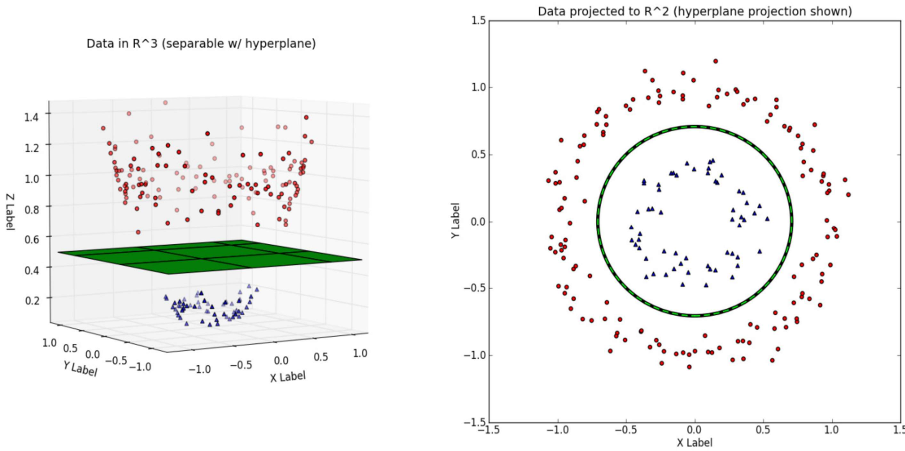
TF-IDF được tính bởi:

$$tf\_idf(t) = tf(t) \times idf(t)$$

2.2. Học máy

Học máy hay máy học (tiếng Anh: Machine learning) là một lĩnh vực của trí tuệ nhân tạo liên quan đến việc nghiên cứu và xây dựng các kĩ thuật cho phép các hệ thống "học" từ dữ liệu để giải quyết những vấn đề cụ thể. Một nhánh của học máy là học sâu phát triển rất mạnh mẽ gần đây và có những kết quả vượt trội so với các phương pháp học máy khác Học máy có liên quan lớn đến thống kê, vì cả hai lĩnh vực đều nghiên cứu việc phân tích dữ liệu, nhưng khác với thống kê, học máy tập trung vào sự phức tạp của các giải thuật trong việc thực thi tính toán.

SVM (Support Vector Machine): Là một thuật toán học máy thuộc loại học có giám sát được sử dụng chủ yếu cho bài toán phân loại và hồi quy. Mục tiêu của SVM là tìm ra một siêu phẳng (hyperplane) trong không gian n chiều (n là số chiều của dữ liệu) sao cho phân cách giữa các lớp là tốt nhất có thể. Đặc điểm của SVM là một trong những phương pháp phân loại mạnh mẽ và phổ biến, có khả năng xử lý cả dữ liệu tuyến tính và phi tuyến tính thông qua sử dụng các hàm kernel [4].



Hình 1: Mô hình đồ thị không gian của SVM

2.3. Tâm lý học

Tâm lý học là ngành khoa học nghiên cứu về tâm trí và hành vi, tìm hiểu về các hiện tượng ý thức và vô thức, cũng như cảm xúc và tư duy. Đây là một bộ môn học thuật với quy mô nghiên cứu rất sâu rộng. Trong nghiên cứu này nhóm tác giả tập trung về các hội chứng tâm lý phổ biến như lo âu, trầm cảm, ám ảnh cưỡng chế, rối loạn lưỡng cực, rối loạn tâm thần, rối loạn tăng động giảm chú ý,… Có một số trạng thái và từ ngữ cụ thể đại diện một phần các vấn đề về tâm lý, bằng cách xác định nhiều các nhóm từ trong những hoàn cảnh cụ thể khác nhau có thể tìm ra được vấn đề tâm lý của người bệnh.

3. Kết quả nghiên cứu và thảo luận

3.1. Dữ liệu thực nghiệm

Dữ liệu thực nghiệm được sử dụng trong nghiên cứu này bao gồm bộ dữ liệu từ những buổi thăm khám tâm lý được bảo mật thông tin và sự hỗ trợ từ một số bác sĩ chuyên khoa tâm lý học. Dữ liệu được kiểm tra và xây dựng bởi những cá nhân có kiến thức liên quan đến tâm lý học.

Dữ liệu được xây dựng dưới dạng các mẫu câu (patterns) và mỗi mẫu câu sẽ được gán 01 tag hợp lý với nội dung của câu đó. Nói cách khác, bộ dữ liệu sẽ có dạng patterns - tag. Ví dụ mẫu “Feeling hopeless” sẽ được gắn tag “Depression”.

Các tag sẽ là các rối loạn tâm lý thường gặp của người mắc bệnh liên quan đến tâm lý. Ví dụ Depression là Hội chứng trầm cảm, ADHD là Hội chứng rối loạn tăng động giảm chú ý.

Dữ liệu được kiểm tra và xây dựng bởi những cá nhân có kiến thức liên quan đến tâm lý học.

Bộ dữ liệu thực nghiệm gồm có 10 tags và 280 patterns cụ thể như Bảng 1.

Bảng 1. Bảng dữ liệu thực nghiệm

Tag	Depr	ADHD	AD	OCD	BD	PTSD	DID	BDD	SUD	Schiz
Input	46	24	24	27	36	28	22	20	17	36

3.2. Mục tiêu thực nghiệm

Mục tiêu thực nghiệm là đánh giá khả năng nhận dạng và cải thiện hệ số tỷ lệ nhận dạng các trạng thái tổn thương tâm lý khác nhau. Đây sẽ là bước đầu tiên của nghiên cứu làm tiền đề cho các bước tiếp theo.

3.3. Kết quả thực nghiệm

3.3.1. Về mô hình học máy

Đầu tiên nhóm tác giả sử dụng thư viện scikit-learn để thực hiện xử lý dữ liệu trước khi đưa vào mô hình học máy, bước này giúp chuẩn bị dữ liệu để có thể sử dụng trong mô hình học máy sau này, bằng cách chuyển đổi văn bản thành dạng số (vectơ TF-IDF) và chuyển đổi nhãn thành dạng số sau đó sử dụng hàm train_test_split từ sklearn.model_selection để chia dữ liệu thành hai tập: Một tập để huấn luyện mô hình và một tập để kiểm tra mô hình.

Trọng số TF-IDF có vai trò cực kỳ quan trọng trong việc phân loại, với trọng số này, một từ có TF-IDF cao trong một văn bản có nghĩa là từ đó xuất hiện nhiều trong văn bản đó hơn so với các văn bản khác, từ đó có thể xác định được mức độ quan trọng của từ đó đối với văn bản.

```
[ ] tfidif_vectorizer = TfidfVectorizer()
    label_encoder = LabelEncoder()
    X = tfidif_vectorizer.fit_transform(X)
    y = label_encoder.fit_transform(y)

[ ] from sklearn.model_selection import train_test_split

    train_sentiment, test_sentiment, train_label, test_label = train_test_split(X, y, test_size = 0.2, shuffle=True, random_state=42)

    print(train_sentiment[0])
    print(train_label[0])

      (0, 268)      0.5149298424993765
      (0, 163)      0.5149298424993765
      (0, 107)      0.2762049952930977
      (0, 141)      0.25308021122566493
      (0, 127)      0.24765701012100866
      (0, 17)       0.29757193703977197
      (0, 136)      0.4236421479179897
    4
```

Hình 2: Tỷ lệ trọng số TF-IDF

Kết quả độ chính xác của mô hình SVM trên dữ liệu kiểm tra là khoảng 48,21 %. Điều này có nghĩa là mô hình đạt được độ chính xác khoảng 48,21 % trong việc dự đoán nhãn cho dữ liệu kiểm tra. Độ chính xác này được coi là tương đối thấp trong một mô hình phân loại văn bản, một trong số các nguyên nhân dẫn đến tỷ lệ này là lượng dữ liệu cho việc huấn luyện vẫn chưa đủ lớn. So sánh với một số mô hình khác kết quả đem lại thấp hơn so với mô hình SVC này tương đối nhiều (khoảng 40 %).

```python
from sklearn.svm import SVC
from sklearn.metrics import accuracy_score


svc_model = SVC(kernel ="linear", C=1.0)
svc_model.fit(train_sentiment, train_label)

y_pred = svc_model.predict(test_sentiment)

accuracy = accuracy_score(test_label, y_pred)
print(accuracy)

0.48214285714285715
```

Hình 3: Tỷ lệ nhận dạng thành công bằng mô hình SVC

Cụ thể thông tin về tỷ lệ nhận dạng thành công theo từng chủ đề (tag) như sau:

Bảng 2. Tỷ lệ nhận dạng thành công theo từng chủ đề

Tag	Depr	ADHD	AD	OCD	BD	PTSD	DID	BDD	SUD	Schiz
Input	46	24	24	27	36	28	22	20	17	36
Correct	21	11	12	14	16	14	12	11	9	15

3.3.2. Về lộ trình điều trị tâm lý

Nhóm đã cơ bản xây dựng hoàn thiện kiến thức cùng quy trình về liệu pháp hành vi và tâm động học. Các liệu pháp và quy trình còn lại sẽ tiếp tục được tìm hiểu và hoàn thiện, sẵn sàng cho một tương lai tới được sử dụng cho chatbot đưa ra các phản hồi và liệu pháp điều trị tốt hơn.

4. Kết luận

Từ bước đầu của nghiên cứu trên, mô hình đề xuất cơ bản đã có thể nhận diện được một phần chính xác các loại tổn thương. Tuy nhiên do hạn chế về kiến thức cũng như nguồn dữ liệu toàn bộ được xây dựng cá nhân, dữ liệu không có sẵn nên chưa thể mang lại hiệu quả tối ưu (nhóm hướng tới tỷ lệ nhận diện đạt 85 %). Nhóm tác giả đề xuất hướng đi sắp tới của nghiên cứu trong tương lai là xây dựng thêm kho dữ liệu tâm lý kết hợp thêm mô hình dữ liệu chatbot PhoBert của tập đoàn Vin để tạo ra các phản hồi có tính mở phục vụ cho mô hình.

TÀI LIỆU THAM KHẢO

[1]. Nguyễn Bá Đạt (2009). *Một số nghiên cứu về sự tổn thương tâm lý*. http://tailieudientu.lrc.tnu.edu.vn/Upload/Collection/brief/brief_28441_32011_21520121433265819209771PB.pdf.

[2]. Nguyễn Tuấn Anh (2019). *Xử lý ngôn ngữ tự nhiên*. https://viblo.asia/p/xu-ly-ngon-ngu-tu-nhien-phan-1-OeVKB8eQlkW.

[3]. Dương Phạm (2016). *TF-IDF*. https://viblo.asia/p/tf-idf-term-frequency-inverse-document-frequency-JQVkVZgKkyd.

[4]. Phạm Văn Toàn (2016). *SVM trong học máy*. https://viblo.asia/p/support-vector-machine-trong-hoc-may-mot-cai-nhin-don-gian-hon-XQZkxoQmewA.

[5]. Jo Hemmings (2020). *How psychology works*. Nxb. Thế giới.

[6]. Shannon Thomas (1988). *Healing the shame that binds you*. Health Communications, Inc.

BBT nhận bài: 14/3/2024; Phản biện xong: 21/3/2024; Chấp nhận đăng: 28/3/2024