

Phân Tích Cảm Xúc Dựa Trên Phương Pháp Học Sâu: Một Số Nghiên Cứu Liên Quan Và So Sánh Đánh Giá

Võ Minh Nguyên¹, Nguyễn Tân Niên¹, Hoàng Đức Dân¹,

Đặng Nhật Huy¹, Trần Phạm Hải Nam¹, Lý Minh Nguyên¹

¹ Khoa Công nghệ Thông tin Kinh doanh, Đại học Kinh tế Thành phố Hồ Chí Minh – Đại học UEH (UEH-HCMC), Việt Nam

Liên hệ tác giả: nguyenvo.31211027655@st.ueh.edu.vn

Tóm tắt

Trong bối cảnh số hóa ngày nay, việc nắm bắt và hiểu cảm xúc của người dùng trên các nền tảng trực tuyến trở nên quan trọng, giúp phân tích hành vi tiêu dùng và quản lý thông tin một cách hiệu quả. Bài báo này tập trung vào việc khám phá và áp dụng các phương pháp học sâu cùng với kỹ thuật xử lý ngôn ngữ tự nhiên (NLP) để phân tích cảm xúc. Về phương pháp thực hiện trong bài báo, chúng tôi đã sử dụng các mô hình học sâu như BERT, BiLSTM và RoBERTa cùng với một loạt các kỹ thuật NLP như phân đoạn từ (Tokenization), gốc từ và dạng từ (Stemming và Lemmatization), biểu diễn từ (Word Embeddings), nhận dạng thực thể có tên và phân tích phụ thuộc câu (Dependency Parsing). Sau đó, áp dụng những phương pháp này vào các tập dữ liệu lớn từ mạng xã hội và các nền tảng thương mại điện tử, chúng tôi đã phân tích cảm xúc của người dùng dựa trên các bình luận và đánh giá về sản phẩm, dịch vụ. Kết quả cho thấy rằng những phương pháp này mang lại kết quả khả quan, giúp đánh giá và hiểu rõ cảm xúc của người dùng một cách tự động và chính xác. Những kết quả này có thể hỗ trợ các doanh nghiệp trong việc tối ưu hóa chiến lược tiếp thị và nâng cao chất lượng sản phẩm dựa trên ý kiến phản hồi từ người tiêu dùng.

Từ khóa: phân tích cảm xúc, phương pháp học sâu, xử lý ngôn ngữ tự nhiên (NLP).

1. Giới thiệu

Ngày nay, dữ liệu từ mạng xã hội và các nền tảng thương mại điện tử đã trở nên vô giá, cung cấp một nguồn thông tin phong phú để phân tích hành vi và tâm lý người tiêu dùng. Việc phân tích cảm xúc trực tuyến không chỉ mang lại cái nhìn chi tiết về thái độ và phản ứng của khách hàng, mà còn là công cụ hữu ích cho các doanh nghiệp trong việc đánh giá và tối ưu hóa các chiến lược tiếp thị của họ, đồng thời nắm bắt và phản hồi nhanh chóng đối với ý kiến của khách hàng. Nghiên cứu của chúng tôi tập trung vào việc ứng dụng các phương pháp học sâu và kỹ thuật xử lý ngôn ngữ tự nhiên (NLP) để phân tích cảm xúc từ dữ liệu trực tuyến. Bài nghiên cứu của chúng tôi được thực hiện với mục tiêu khám phá cách thức áp dụng các mô hình học sâu để hiểu và dự đoán cảm xúc của người dùng một cách chính xác. Đồng thời, chúng tôi cũng đánh giá hiệu quả của nhiều phương pháp và mô hình khác nhau trong lĩnh vực này.

Trong bối cảnh này, phương pháp học sâu đã nổi lên như một công cụ mạnh mẽ trong phân tích cảm xúc trực tuyến. Sự tiến bộ trong lĩnh vực học máy đã tạo ra các mô hình như BERT (Bidirectional Encoder Representations from Transformers), BiLSTM (Bidirectional Long Short-Term Memory), và RoBERTa (A

Robustly Optimized BERT Pretraining Approach), đều góp phần làm tăng khả năng hiểu và xử lý ngôn ngữ tự nhiên một cách chính xác. Các mô hình này có thể phân tích mối quan hệ giữa từ ngữ và ngữ cảnh một cách phức tạp, từ đó cải thiện đáng kể khả năng nhận diện và dự đoán cảm xúc của người dùng. Bên cạnh đó, kỹ thuật NLP cũng đóng một vai trò không thể thiếu, giúp biểu diễn ngôn ngữ một cách hiệu quả và dễ dàng tích hợp vào các mô hình học sâu.

Nghiên cứu trong lĩnh vực phân tích cảm xúc trực tuyến đã đạt được những tiến bộ đáng kể, đặc biệt là thông qua việc áp dụng phương pháp học sâu và xử lý ngôn ngữ tự nhiên (NLP). Các công trình nghiên cứu tiêu biểu như của (Socher và cộng sự, 2013) đã phát triển mô hình Recursive Neural Tensor Network (RNTN), một bước đột phá trong việc phân tích cảm xúc từ văn bản phức tạp. Hay (Yoon & Kim, 2017), cũng đã góp phần vào lĩnh vực này bằng cách sử dụng mạng nơ-ron hồi quy trong việc phân tích cảm xúc từ văn bản. Gần đây, sự chú ý đã được chuyển hướng tới việc tích hợp các mô hình học sâu tiên tiến như BERT và các biến thể của nó, với (Delgadillo và cộng sự, 2023; Devlin và cộng sự, 2018) đã giới thiệu BERT, một mô hình ngôn ngữ mạnh mẽ được huấn luyện trên một lượng lớn dữ liệu từ internet, cho thấy hiệu suất vượt trội so với các phương pháp trước đây trong nhiều tác vụ NLP, bao gồm phân tích cảm xúc. Các nghiên cứu này đã cung cấp một nền tảng vững chắc cho các nghiên cứu tiếp theo và đã thúc đẩy sự phát triển của lĩnh vực này.

Nhìn chung, việc kết hợp giữa học sâu và các kỹ thuật xử lý ngôn ngữ tự nhiên (NLP) đang mở ra tiềm năng lớn cho việc phân tích cảm xúc trực tuyến. Qua bài nghiên cứu này, chúng tôi mong muốn đóng góp vào lĩnh vực này bằng việc cung cấp cái nhìn sâu sắc hơn và phân tích so sánh chi tiết về hiệu suất của các phương pháp và mô hình khác nhau. Chúng tôi hy vọng rằng, thông qua việc này, bài nghiên cứu sẽ góp phần vào việc phát triển các phương pháp phân tích cảm xúc trực tuyến, giúp cải thiện sự hiểu biết về cảm xúc của người dùng và tăng cường khả năng dự đoán trong quản lý doanh nghiệp và tương tác với khách hàng.

Để cập nhật và phản ánh rõ về sự tiến bộ trong lĩnh vực này, chúng tôi sẽ tiếp tục giới thiệu một số nghiên cứu mới nhất liên quan đến chủ đề của bài báo. Những nghiên cứu này không chỉ mở rộng kiến thức mà còn cung cấp cơ sở lý thuyết và thực tế cho việc phát triển nghiên cứu của chúng tôi. Cụ thể, trong nghiên cứu của (Batra và cộng sự, 2021) và (Delgadillo và cộng sự, 2023), các tác giả đã sử dụng mô hình BERT để phân tích cảm xúc trong các bài viết trên mạng xã hội và mạng xã hội tài chính. Kết quả cho thấy hiệu suất của mô hình BERT được đánh giá cao hơn so với các phương pháp truyền thống. Kết quả cho thấy rằng BERT có khả năng nhận diện cảm xúc của người dùng với độ chính xác cao hơn so với các phương pháp truyền thống khác. Hay trong nghiên cứu của (Xu và cộng sự, 2019), tác giả đã phát triển một mô hình ERNIE-BiLSTM (Enhanced Representation through kNowledge Integration – Bidirectional Long Short-Term Memory) để nhận diện cảm xúc trong các bình luận trực tuyến. Kết quả thử nghiệm đã cho thấy rằng mô hình đề xuất có hiệu suất tốt và ổn định, đồng thời vượt qua nhiều phương pháp tiên tiến khác. Ngoài ra, nghiên cứu của (Chamekh và cộng sự, 2022) và (Yi Liu và cộng sự, 2020) đã áp dụng phương pháp học sâu vào phân tích cảm xúc trong các đánh giá sản phẩm trên các sàn thương mại điện tử. Kết quả của họ đã chỉ ra rằng việc sử dụng mô hình học sâu có thể cải thiện đáng kể hiệu suất của việc phân tích cảm xúc so với các phương pháp truyền thống. Những nghiên cứu này cùng với các nghiên cứu khác đã tiếp tục mở ra nhiều hướng tiếp cận và cơ hội mới trong lĩnh vực phân tích cảm xúc dựa trên phương pháp học sâu và NLP.

Chúng tôi sẽ tiếp tục tìm hiểu và áp dụng những tiến bộ này vào nghiên cứu của mình, nhằm mang lại những đóng góp mới và cải thiện hiệu suất của phương pháp phân tích cảm xúc trực tuyến.

Trên cơ sở những nghiên cứu và phát triển mới nhất trong lĩnh vực này, chúng tôi hy vọng rằng bài báo này sẽ cung cấp cái nhìn toàn diện và chi tiết về phương pháp, mô hình và kỹ thuật hiện đại nhất trong phân tích cảm xúc dựa trên học sâu, đồng thời mở ra những hướng nghiên cứu tiềm năng cho cộng đồng nghiên cứu trong tương lai.

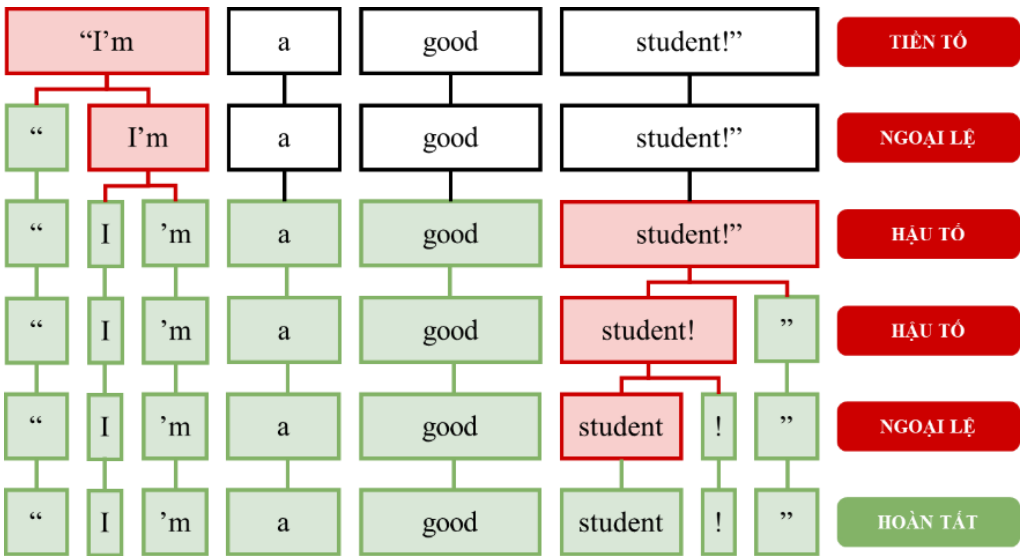
2. Nền tảng lý thuyết

2.1. Các kỹ thuật xử lý văn bản

Trong lĩnh vực phân tích cảm xúc, việc xử lý văn bản đóng vai trò quan trọng không thể phủ nhận. Đặc biệt, khi dữ liệu đầu vào chủ yếu là các đoạn văn bản từ mạng xã hội và các trang thương mại điện tử, nó trở nên phức tạp hơn với cấu trúc đa dạng và khác biệt so với những mẫu câu tiêu chuẩn thường được sử dụng. Một trong những thách thức lớn nhất của việc xử lý văn bản trong phân tích cảm xúc là sự đa dạng về ngôn ngữ, cấu trúc câu, từ vựng và ngữ cảnh. Các đoạn văn bản từ mạng xã hội thường chứa các từ ngữ viết tắt, từ lóng, ngôn ngữ không chuẩn và thậm chí là các biểu cảm cảm xúc khó đánh giá. Điều này tạo ra một thách thức lớn trong việc hiểu đúng nghĩa của các từ và câu trong ngữ cảnh của chúng. Trong bối cảnh này, việc áp dụng các kỹ thuật xử lý văn bản trở thành bước quan trọng để chuẩn hóa và tiền xử lý dữ liệu, tạo điều kiện thuận lợi cho quá trình phân tích cảm xúc. Một số kỹ thuật NLP sẽ được thảo luận trong phần này.

2.1.1. Tokenization

Tokenization là một bước quan trọng trong xử lý văn bản trong lĩnh vực xử lý ngôn ngữ tự nhiên (NLP). Quá trình này nhằm chia các đoạn văn bản thành các đơn vị nhỏ hơn gọi là "token", có thể là từ, cụm từ hoặc ký tự. Mục tiêu của tokenization là chuẩn hóa và chia nhỏ dữ liệu văn bản để dễ dàng xử lý và phân tích. Được minh họa cụ thể như **Hình 1**.



Hình 1. Phương pháp “token” hóa trong NLP.

Trong nghiên cứu của (Sanderson, 2010), tác giả đã áp dụng các kỹ thuật tokenization hiệu quả để xử lý văn bản tự nhiên, bao gồm cả việc sử dụng biểu thức chính quy và các quy tắc ngữ cảnh. Kết quả của họ cho thấy rằng các phương pháp tokenization này đã đạt được hiệu suất cao trong việc phân tách và chuẩn hóa dữ liệu, bất kể cấu trúc ngôn ngữ đa dạng và khác biệt. Trong một nghiên cứu khác của (Bird và cộng sự, 2009), tác giả đã mô tả các kỹ thuật tokenization phức tạp hơn, bao gồm cả việc sử dụng các biểu thức chính quy và thuật toán phân tích ngữ cảnh để xác định các ranh giới token một cách chính xác hơn. Kết quả của họ đã chỉ ra rằng các phương pháp tokenization này có khả năng xử lý các cấu trúc ngôn ngữ phức tạp và đa dạng một cách hiệu quả, giúp cải thiện hiệu suất của quá trình phân tích cảm xúc và xử lý ngôn ngữ tự nhiên.

Tóm lại, các nghiên cứu trên đã chứng minh rằng các kỹ thuật tokenization có thể khắc phục được sự phức tạp về cấu trúc đa dạng và khác biệt trong các đoạn văn bản từ mạng xã hội và các trang thương mại điện tử. Các phương pháp này đóng vai trò quan trọng trong việc chuẩn hóa và chia nhỏ dữ liệu văn bản, từ đó giúp cải thiện hiệu suất của quá trình phân tích cảm xúc và xử lý ngôn ngữ tự nhiên trong lĩnh vực NLP.

2.1.2. Stemming và Lemmatization

Stemming và Lemmatization là hai kỹ thuật quan trọng trong xử lý ngôn ngữ tự nhiên (NLP) nhằm chuẩn hóa các từ về dạng gốc của chúng, từ đó giảm thiểu sự biến thể và tăng khả năng đối chiếu giữa các từ tương đồng. Mặc dù cả hai kỹ thuật đều nhằm mục đích tương tự, nhưng cách thực hiện và hiệu suất của chúng có thể khác nhau.

Stemming là quá trình cắt bớt các phần cuối của từ để đạt được dạng gốc hoặc hậu tố. Mục tiêu của Stemming là loại bỏ các hậu tố không cần thiết để tạo ra một từ gốc phù hợp, giúp giảm thiểu sự biến thể và tăng cường khả năng đối chiếu giữa các từ tương đồng. Một trong những thuật toán Stemming phổ biến nhất là thuật toán Porter, được phát triển bởi Martin Porter vào những năm 1980.

Lemmatization, mặt khác, là quá trình chuyển đổi từ về dạng gốc sử dụng từ điển hoặc ngữ pháp. So với Stemming, Lemmatization cung cấp kết quả chuẩn xác hơn bởi vì nó dựa vào thông tin từ điển để xác định từ gốc của một từ. Mặc dù Lemmatization có thể tốn nhiều thời gian hơn và đòi hỏi tài nguyên tính toán lớn hơn, nhưng nó thường cho kết quả chính xác hơn trong các trường hợp phức tạp.

Các nghiên cứu gần đây đã nhấn mạnh vai trò quan trọng của Stemming và Lemmatization trong phân tích cảm xúc. (B. Liu & Zhang, 2012) đã áp dụng Stemming để giảm biến thể từ trong quá trình phân tích cảm xúc từ dữ liệu ngôn ngữ tự nhiên. Kết quả cho thấy rằng việc này đã cải thiện hiệu suất của mô hình phân tích cảm xúc dựa trên học sâu. Một nghiên cứu khác của (L. Zhang và cộng sự, 2018) tiến xa hơn bằng cách kết hợp cả Stemming và Lemmatization trong tiền xử lý dữ liệu cho ứng dụng phân tích cảm xúc trực tuyến. Kết quả của họ chứng minh rằng sự kết hợp này đã đóng góp đáng kể vào độ chính xác và hiệu suất của mô hình phân tích cảm xúc. Ngoài ra, các nghiên cứu trước đó của (Lovins, 1968) và (Scherer, 1984) cũng đã đề cập đến vai trò quan trọng của Stemming và Lemmatization trong xử lý ngôn ngữ tự nhiên và phân tích cảm xúc.

Việc sử dụng Stemming và Lemmatization trong phân tích cảm xúc, cũng như trong xử lý các câu văn trên các trang mạng xã hội và sàn thương mại điện tử, đã được nhiều nghiên cứu khẳng định là mang lại lợi ích

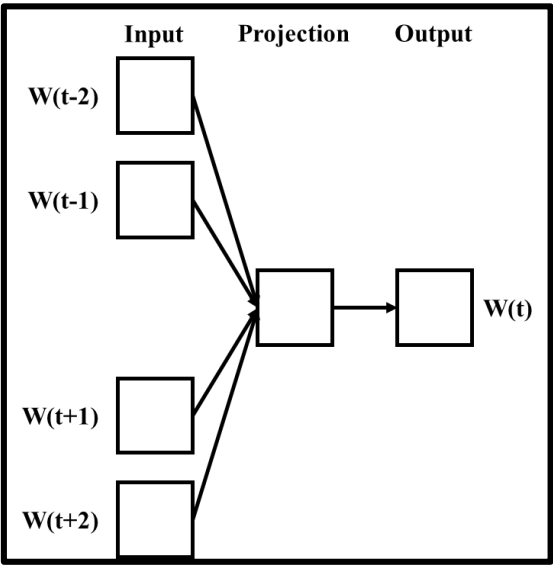
đáng kể. Những kỹ thuật này giúp chuẩn hóa dữ liệu, giảm biến thể từ và tăng cường khả năng so sánh giữa các từ tương đồng, từ đó cải thiện hiệu suất và độ chính xác của quá trình phân tích cảm xúc.

2.1.3. Word Embeddings

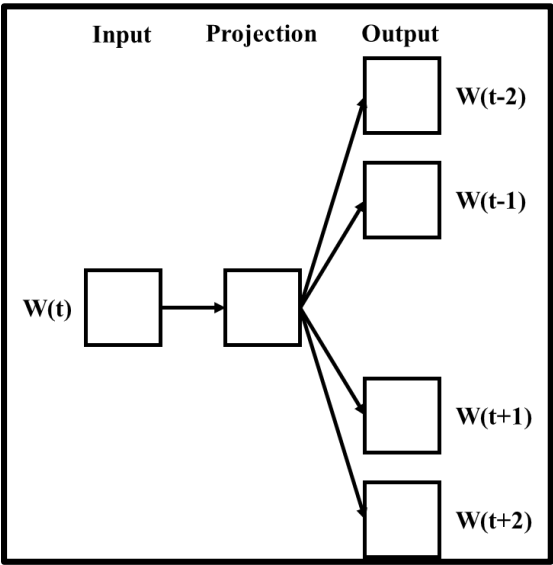
Trong lĩnh vực xử lý ngôn ngữ tự nhiên (NLP), Word Embeddings đóng vai trò quan trọng trong việc biểu diễn từ dưới dạng các vector số học trong không gian đa chiều. Điều này không chỉ giúp mô hình học sâu hiểu được ngữ cảnh và mối quan hệ giữa các từ, mà còn tạo ra nhiều lợi ích khác.

Một trong những lợi ích chính của Word Embeddings là khả năng cải thiện hiểu biết về ngữ cảnh xung quanh từ. Điều này giúp mô hình phân tích cảm xúc và nhận biết các từ tương tự một cách chính xác hơn. Ngoài ra, việc biểu diễn từ dưới dạng vector số cũng giúp mô hình học sâu hiểu được mối quan hệ phức tạp giữa các từ. Điều này không chỉ cải thiện độ chính xác trong việc phân tích cảm xúc, mà còn tăng cường khả năng áp dụng của mô hình trong các ứng dụng thực tế. Như vậy, Word Embeddings không chỉ cải thiện hiệu suất của các mô hình học sâu trong NLP, mà còn mở rộng khả năng áp dụng của chúng trong thực tế.

Công trình nghiên cứu của (Tomáš Mikolov và cộng sự, 2013) đã đánh dấu một bước tiến quan trọng trong lĩnh vực xử lý ngôn ngữ tự nhiên (NLP). Tác giả đã giới thiệu hai mô hình mới là Continuous Bag of Words (CBOW) và Skip-gram, được minh họa qua **Hình 2** và **Hình 3**. Những mô hình này cung cấp cách biểu diễn từ vựng hiệu quả, giúp các mô hình học sâu có khả năng hiểu và phân tích ngôn ngữ một cách toàn diện hơn.



Hình 2. CBOW.

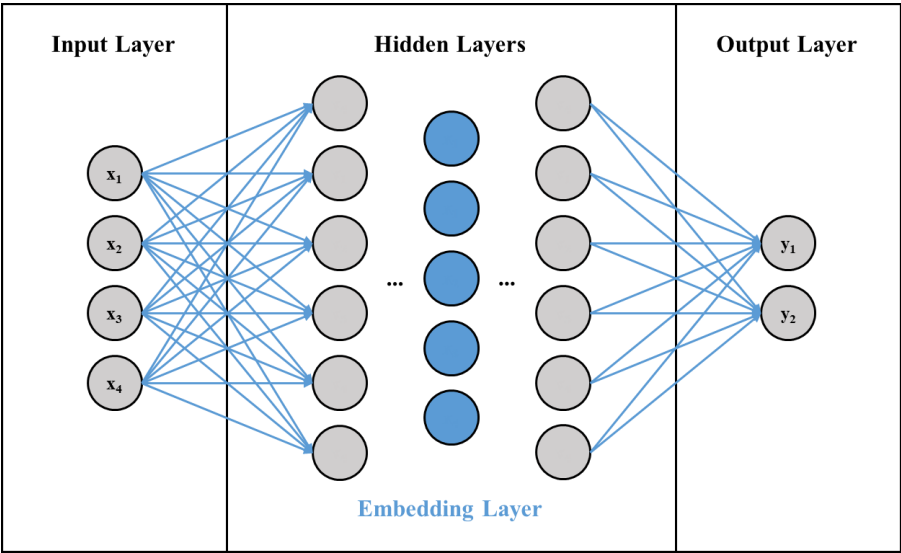


Hình 3. Skip-gram.

Ngoài ra, (Tomas Mikolov, Kai Chen, và cộng sự, 2013) cũng đã phát triển phương pháp Word2Vec được đề cập trong một công trình khác. Phương pháp này, mặc dù phổ biến, nhưng lại mang lại hiệu quả đáng kể. Word2Vec không chỉ giúp mô hình học sâu nắm bắt được ngữ cảnh và mối quan hệ giữa các từ, mà còn cho phép biểu diễn từ dưới dạng vector số học. Điều này mở ra khả năng mới cho các mô hình trong việc hiểu cách sử dụng từ trong nhiều ngữ cảnh khác nhau và mối quan hệ giữa chúng, từ đó cải thiện đáng kể khả năng phân tích ngôn ngữ tự nhiên và tạo tiền đề cho việc áp dụng Word2Vec trong nhiều ứng dụng thực tế.

Bên cạnh đó, nghiên cứu của (Pennington và cộng sự, 2014) đã đưa ra góc nhìn mới trong lĩnh vực Word Embeddings. Tác giả không chỉ mở rộng khả năng của Word Embeddings thông qua việc tích hợp chúng

với mạng Neural Network, mà còn chứng minh được rằng sự kết hợp này có thể nâng cao hiệu suất của các ứng dụng NLP, đặc biệt là trong lĩnh vực phân tích cảm xúc. Điều này được minh họa qua **Hình 4**, nơi các từ được biểu diễn một cách phức tạp hơn. Ngoài ra, công trình của (Tomas Mikolov, Ilya Sutskever, và cộng sự, 2013) cũng góp phần khẳng định vị thế của Word Embeddings trong ngành NLP. Họ đã chỉ ra rằng, việc áp dụng Word Embeddings không chỉ đơn thuần là một phương pháp mới mẻ, mà còn là một cách tiếp cận hiệu quả để cải thiện hiệu suất của các mô hình học sâu. Những phát hiện này mở ra hướng đi mới cho việc phát triển các ứng dụng NLP, đồng thời củng cố thêm bằng chứng cho tiềm năng của Word Embeddings trong việc xử lý ngôn ngữ tự nhiên.



Hình 4. Các tầng Word Embeddings trong mạng nơ-ron.

Nhìn chung, Word Embeddings là một công cụ mạnh mẽ trong xử lý ngôn ngữ tự nhiên, đặc biệt trong việc phân tích cảm xúc dựa trên học sâu và xử lý các câu văn trên mạng xã hội hoặc sàn thương mại điện tử. Sử dụng Word Embeddings giúp cải thiện khả năng hiểu và phân tích ngôn ngữ tự nhiên của mô hình, nâng cao độ chính xác và hiệu suất của các ứng dụng NLP.

2.1.4. *Named Entity Recognition (NER)*

Named Entity Recognition (NER), là một quy trình tự động nhằm xác định và phân loại các thực thể có tên trong văn bản, như tên riêng, địa danh, tên tổ chức, ngày tháng, số lượng và các yếu tố khác có ý nghĩa trong ngữ cảnh của văn bản. Mục tiêu của NER là làm cho việc nhận diện các thực thể này trở nên tự động và chính xác, giúp tạo ra cấu trúc dữ liệu có tổ chức hơn từ các đoạn văn bản không cấu trúc. Quá trình này thường được thực hiện bằng cách sử dụng các phương pháp và thuật toán máy học, trong đó có mạng nơ-ron học sâu (Deep Learning), Support Vector Machines (SVM), Hidden Markov Models (HMM), và Conditional Random Fields (CRF). NER có ứng dụng rộng rãi trong nhiều lĩnh vực, bao gồm tổ chức tin tức, phân tích cảm xúc, dịch thuật máy, tổ chức tri thức, trích xuất thông tin, và nhiều ứng dụng khác liên quan đến xử lý văn bản và ngôn ngữ tự nhiên.

Trong lĩnh vực phân tích cảm xúc dựa trên học sâu và xử lý văn bản trên mạng xã hội hoặc sàn thương mại điện tử, việc sử dụng Named Entity Recognition (NER) mang lại nhiều lợi ích đáng kể. Đầu tiên, NER giúp mô hình hiểu rõ hơn ngữ cảnh cụ thể của văn bản bằng cách nhận diện và phân loại các thực thể có tên. Điều

này đóng vai trò quan trọng trong việc cải thiện khả năng phân tích cảm xúc bằng cách tập trung vào các phần quan trọng của văn bản. Thứ hai, việc nhận diện và phân loại các thực thể có tên giúp tăng độ chính xác trong phân tích cảm xúc bằng cách chỉ ra những phần quan trọng của văn bản cần được xem xét. Điều này không chỉ giúp cải thiện hiệu suất của các mô hình học sâu mà còn giúp trích xuất thông tin chính xác từ văn bản, đặc biệt là trong việc phân tích cảm xúc của người dùng trên mạng xã hội và các sàn thương mại điện tử.

Các nghiên cứu đã chứng minh rằng việc sử dụng Named Entity Recognition (NER) mang lại nhiều lợi ích trong việc phân tích cảm xúc và xử lý văn bản. Ví dụ, nghiên cứu của (Lample và cộng sự, 2016) đã thành công trong việc phân tích cảm xúc bằng cách xác định các thực thể liên quan đến ngữ cảnh cụ thể, trong khi nghiên cứu của (Xie và cộng sự, 2018) đã chỉ ra rằng việc sử dụng NER là một phương pháp hiệu quả để cải thiện độ chính xác và hiệu suất của các mô hình học sâu.

Về cơ bản, Named Entity Recognition (NER) đóng vai trò không thể phủ nhận trong việc tiến xa hơn trong lĩnh vực NLP, đặc biệt là trong phân tích cảm xúc. Bằng cách nhận diện và phân loại các thực thể có tên, NER không chỉ giúp cải thiện khả năng hiểu và phân tích ngữ cảnh cụ thể trong văn bản mà còn nâng cao độ chính xác và hiệu suất của các mô hình NLP. Sự tiến bộ trong NER đã mở ra nhiều cơ hội mới để ứng dụng trong nhiều lĩnh vực, từ phân tích dữ liệu đến dịch thuật tự động và quảng cáo trực tuyến. Với tiềm năng không ngừng của công nghệ này, ta có thể kỳ vọng vào sự phát triển mạnh mẽ của NLP trong tương lai, mang lại những tiện ích và ứng dụng tiên tiến hơn cho xã hội và doanh nghiệp.

2.1.5. Dependency Parsing

Dependency Parsing là một công cụ quan trọng trong việc phân tích cú pháp của câu văn bản, giúp xác định mối quan hệ phụ thuộc giữa các từ. Thông qua việc phân tích cú pháp, Dependency Parsing giúp giải thích cách các từ trong câu được kết nối với nhau và mối quan hệ logic giữa chúng. Bằng cách này, Dependency Parsing không chỉ đơn giản là xác định cấu trúc của câu mà còn làm nổi bật những mối liên kết ý nghĩa giữa các thành phần ngôn ngữ.

Trong lĩnh vực phân tích cảm xúc và xử lý văn bản, Dependency Parsing trở thành một phần không thể thiếu. Việc hiểu được cách các từ được sắp xếp và liên kết với nhau trong câu giúp mô hình hiểu được ngữ cảnh và ý nghĩa của thông điệp. Điều này làm cho việc phân tích cảm xúc và trích xuất thông tin từ văn bản trở nên hiệu quả hơn. Ngoài ra, Dependency Parsing còn cải thiện khả năng xử lý ngôn ngữ tự nhiên bằng cách hiểu được cấu trúc và ý nghĩa của các câu văn. Điều này có thể hỗ trợ trong việc phân tích và xử lý các câu văn trên các trang mạng xã hội và sàn thương mại điện tử một cách hiệu quả hơn. Công cụ này giúp mô hình tạo ra các phân tích và trích xuất thông tin chính xác từ văn bản, tăng cường khả năng hiểu và tương tác với ngôn ngữ tự nhiên.

Một số công trình nghiên cứu liên quan đã chứng minh sự quan trọng của Dependency Parsing trong việc cải thiện khả năng phân tích cảm xúc và xử lý văn bản trong lĩnh vực xử lý ngôn ngữ tự nhiên (NLP). Ví dụ, nghiên cứu của (D. Chen & Manning, 2014) là một trong những công trình nổi bật, khẳng định rằng Dependency Parsing không chỉ là một phần không thể thiếu của NLP mà còn đã được áp dụng rộng rãi trong các ứng dụng phân tích cảm xúc. Qua việc hiểu rõ hơn về ngữ cảnh và mối quan hệ giữa các từ trong văn

bản, Dependency Parsing đóng vai trò quan trọng trong việc phát hiện các yếu tố quan trọng và tập trung vào phần quan trọng nhất của văn bản để phân tích cảm xúc một cách chính xác.

Đồng thời, nghiên cứu của (D. Chen & Manning, 2014) cũng nhấn mạnh rằng việc sử dụng Dependency Parsing giúp cải thiện khả năng xử lý ngôn ngữ tự nhiên bằng cách hiểu được cấu trúc và ý nghĩa của các câu văn. Điều này không chỉ hỗ trợ trong việc phân tích và xử lý các câu văn trên các trang mạng xã hội và sàn thương mại điện tử một cách hiệu quả hơn mà còn nâng cao độ chính xác và hiệu suất của các mô hình NLP.

Ngoài ra, nghiên cứu của (Socher và cộng sự, 2013) cũng đã cung cấp bằng chứng rõ ràng về tính hiệu quả của Dependency Parsing trong việc cải thiện độ chính xác và hiệu suất của các mô hình học sâu trong phân tích cảm xúc trên mạng xã hội. Sự kết hợp giữa Dependency Parsing và các phương pháp học sâu đã tạo ra những tiến bộ đáng kể trong việc hiểu và phân tích ngôn ngữ tự nhiên, từ đó đưa ra những dự đoán chính xác về ý nghĩa và nội dung của văn bản.

Tổng quan lại, sự kết hợp giữa Dependency Parsing và các mô hình học sâu không chỉ làm tăng cường hiệu quả của phân tích cảm xúc và xử lý văn bản mà còn mở ra những tiềm năng mới trong việc hiểu và tận dụng ngôn ngữ tự nhiên để áp dụng vào nhiều lĩnh vực thực tiễn khác nhau.

Trong khía cạnh NLP, Dependency Parsing đóng vai trò quan trọng bằng cách cung cấp cái nhìn chi tiết và cấu trúc về cách các từ liên kết với nhau trong một câu. Điều này không chỉ giúp mô hình hiểu rõ hơn về ngữ cảnh và mối quan hệ giữa các từ mà còn cải thiện đáng kể khả năng phân tích cảm xúc và xử lý văn bản. Khi một mô hình học sâu được trang bị Dependency Parsing, nó có thể chính xác xác định các yếu tố quan trọng trong câu văn và tập trung vào phần quan trọng nhất để phân tích cảm xúc, từ đó nâng cao độ chính xác và hiệu suất của mô hình. Đồng thời, khả năng hiểu được cấu trúc câu giúp mô hình tạo ra các dự đoán chính xác hơn về ý nghĩa và nội dung của văn bản. Do đó, Dependency Parsing không chỉ là một công cụ hữu ích mà còn là một phần quan trọng trong việc đẩy mạnh khả năng của các mô hình học sâu trong NLP.

2.2. *Mô hình học sâu*

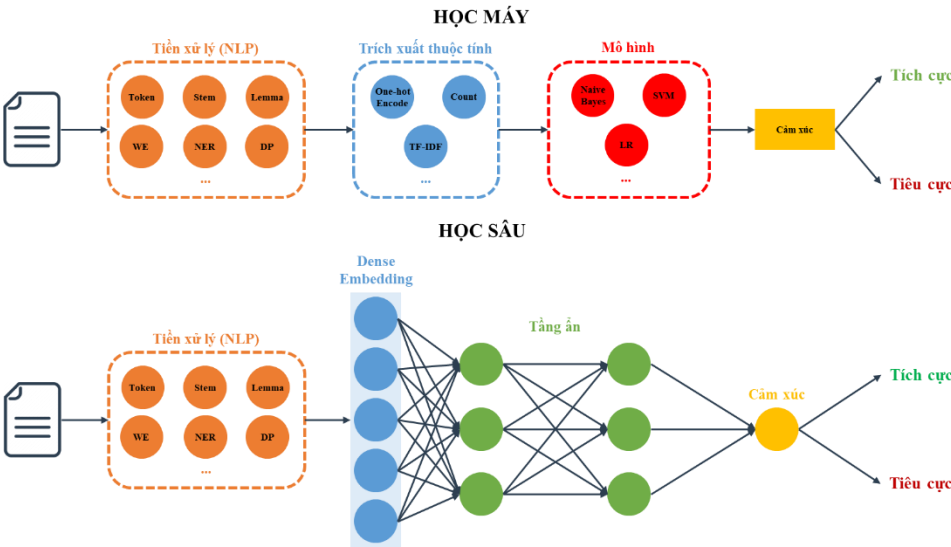
Trong thập kỷ gần đây, mô hình học sâu đã trở thành công cụ quan trọng trong nhiều lĩnh vực của khoa học máy và xử lý ngôn ngữ tự nhiên (NLP), bao gồm cả phân tích cảm xúc. Mô hình học sâu là một dạng của mạng nơ-ron sâu, có khả năng học các biểu diễn phức tạp từ dữ liệu thông qua việc lặp đi lặp lại của các lớp. Sự mạnh mẽ của mô hình học sâu đến từ khả năng tự học và tự điều chỉnh dựa trên dữ liệu, giúp chúng có thể hiểu và mô hình hóa các mối quan hệ phức tạp giữa các đặc trưng của dữ liệu.

Một trong những ưu điểm chính của mô hình học sâu là khả năng học được các biểu diễn tự nhiên của dữ liệu. Trong phân tích cảm xúc, điều này có ý nghĩa quan trọng vì ngôn ngữ tự nhiên thường chứa các mối quan hệ phức tạp và không gian đa chiều của các từ và cụm từ. Các mô hình học sâu có khả năng học và biểu diễn các biểu đạt ngôn ngữ này một cách hiệu quả, giúp họ có thể nhận dạng và hiểu các biểu diễn ngôn ngữ tự nhiên trong các văn bản.

Một số nghiên cứu đã chứng minh rằng mô hình học sâu mang lại hiệu suất tốt hơn so với các mô hình học máy truyền thống trong việc phân tích cảm xúc. Ví dụ, trong nghiên cứu của (Li và cộng sự, 2022), các tác giả đã so sánh hiệu suất của một mô hình BiLSTM-CRF với các phương pháp truyền thống khác trong việc

nhận diện cảm xúc trong các bình luận trực tuyến. Kết quả cho thấy rằng mô hình học sâu này vượt trội về độ chính xác và độ ổn định, đặc biệt là trong việc xử lý các ngữ cảnh phức tạp.

Hay, theo nghiên cứu của (Dang và cộng sự, 2020), cả hai phương pháp học máy truyền thống và học sâu đều được áp dụng trong việc phân loại cảm xúc. Tuy nhiên, chúng có những khác biệt đáng chú ý. Phương pháp học máy truyền thống yêu cầu việc xác định và trích xuất các tính năng một cách thủ công hoặc thông qua các phương pháp lựa chọn tính năng. Các mô hình phân loại như SVM, mạng Bayes và cây quyết định đều sử dụng phương pháp này. Trái lại, học sâu cho phép việc học và trích xuất các tính năng một cách tự động, dẫn đến độ chính xác và hiệu suất tốt hơn. Thậm chí, các siêu tham số của mô hình phân loại cũng được đo lường một cách tự động. Ở **Hình 5** minh họa sự khác biệt giữa hai phương pháp này.



Hình 5. So sánh hai phương pháp phân loại cảm xúc: Học máy và Học sâu.

Ngoài ra, một số nghiên cứu khác cũng chứng minh rằng việc kết hợp các mô hình học sâu với các kỹ thuật xử lý ngôn ngữ tự nhiên (NLP) có thể mang lại hiệu suất tốt hơn. Trong nghiên cứu của (Acheampong và cộng sự, 2021), các tác giả đã sử dụng kết hợp của mô hình BERT và các phương pháp xử lý văn bản NLP để phân tích cảm xúc trong các bài viết trên mạng xã hội. Kết quả cho thấy rằng phương pháp này đạt được hiệu suất tốt nhất so với các phương pháp khác, nhấn mạnh sức mạnh của việc kết hợp giữa học sâu và xử lý ngôn ngữ tự nhiên trong phân tích cảm xúc.

Như vậy, mô hình học sâu không chỉ có khả năng học được các biểu diễn phức tạp của dữ liệu mà còn mang lại hiệu suất tốt hơn trong việc phân tích cảm xúc so với các phương pháp truyền thống. Sự kết hợp giữa mô hình học sâu và các kỹ thuật xử lý ngôn ngữ tự nhiên là một hướng tiếp cận promising và có thể mở ra nhiều cơ hội mới trong lĩnh vực phân tích cảm xúc. Một số kỹ thuật học sâu sẽ được thảo luận trong phần này.

2.2.1. BERT (Bidirectional Encoder Representations from Transformers)

Kiến trúc của BERT được thiết kế một cách tinh vi, kết hợp các thành phần chủ chốt của Transformer cùng với một số cải tiến độc đáo, nhằm tối ưu hóa cho các mục tiêu cụ thể của mô hình. Dưới đây là những yếu tố nổi bật trong kiến trúc của BERT.

Cơ sở của Transformer (Transformer Base)

BERT được xây dựng dựa trên kiến trúc Transformer, một mô hình mạng nơ-ron sâu (DNN) đặc biệt được tối ưu hóa cho việc xử lý dữ liệu tuần tự như văn bản. Kiến trúc Transformer bao gồm các lớp mã hóa và giải mã, trong đó mỗi lớp chứa một số lớp tự chú ý và lớp mạng nơ-ron truyền thẳng. Đặc điểm nổi bật của Transformer là khả năng mã hóa ngữ cảnh toàn diện từ cả hai phía của câu, điều này đã góp phần cải thiện hiệu suất trong việc hiểu ngôn ngữ tự nhiên.

Mã hóa hai chiều

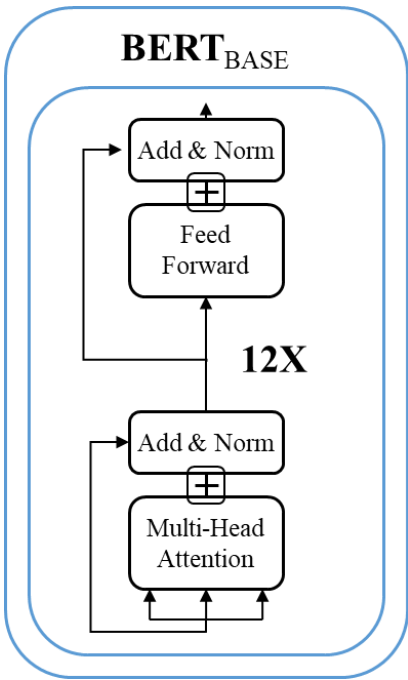
BERT tinh giản kiến trúc của Transformer bằng cách loại bỏ lớp giải mã và chỉ giữ lại phần mã hóa. Điều này cho phép BERT tập trung vào việc hiểu ngữ cảnh từ cả hai phía của câu, thay vì chỉ từ trái qua phải như trong các mô hình truyền thống. Sự thay đổi này giúp BERT hiểu văn bản một cách toàn diện hơn, từ đó cải thiện hiệu suất cho nhiều nhiệm vụ xử lý ngôn ngữ tự nhiên (NLP).

Mô hình học không giám sát (Unsupervised Learning)

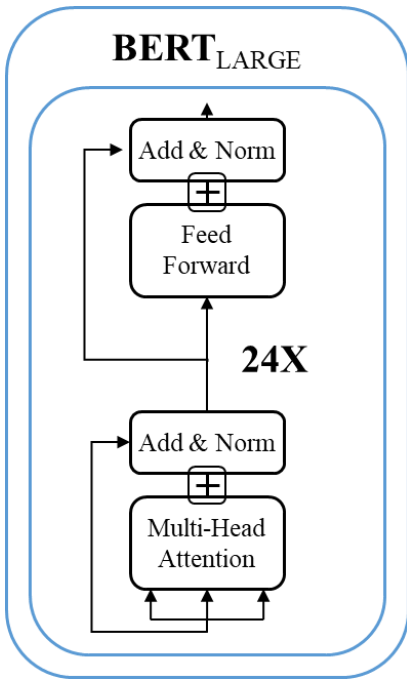
BERT được huấn luyện trên dữ liệu lớn với một nhiệm vụ không giám sát gọi là "Masked Language Model" (MLM) và "Next Sentence Prediction" (NSP). MLM yêu cầu mô hình dự đoán các từ bị che đi trong câu, trong khi NSP đánh giá xem hai câu có liên quan logic hay không. Sự kết hợp của hai nhiệm vụ này giúp BERT hiểu ngữ cảnh của từng từ trong câu và cách các câu liên quan đến nhau.

Tính khả dụng và hiệu suất

BERT có nhiều phiên bản, từ BERT Base, BERT Large cho đến BERT Tiny, tùy thuộc vào kích thước và hiệu suất mong muốn, được minh họa qua **Hình 6** và **Hình 7**. Điều này tạo điều kiện cho các nhà nghiên cứu và nhà phát triển có thể lựa chọn mô hình phù hợp với yêu cầu cụ thể của họ, từ việc phát triển ứng dụng với tài nguyên hạn chế đến việc nghiên cứu với dữ liệu lớn.



Hình 6. 110M Thông số.



Hình 7. 340M Thông số.

Nhìn chung, kiến trúc BERT kết hợp những yếu tố quan trọng nhất của Transformer với các cải tiến đáng kể để tạo ra một mô hình học sâu vượt trội trong việc hiểu và xử lý ngôn ngữ tự nhiên. Một trong những điểm mạnh của BERT là khả năng biểu diễn ngôn ngữ tự nhiên một cách phong phú và đa dạng. Do đó, BERT có thể áp dụng vào nhiều nhiệm vụ NLP khác nhau, từ phân loại văn bản, dự đoán từ tiếp theo cho đến phân tích cảm xúc.

Việc áp dụng của BERT trong lĩnh vực phân tích cảm xúc mang lại nhiều lợi ích đáng chú ý. Nhờ khả năng hiểu và mô hình hóa ngôn ngữ tự nhiên một cách hiệu quả, BERT có thể nhận diện và phân tích cảm xúc của người dùng từ các đoạn văn bản một cách chính xác và toàn diện. Đặc biệt, BERT có khả năng xử lý cả ngữ cảnh phức tạp và từ ngữ đa nghĩa, giúp tăng độ chính xác và độ tin cậy của phân tích cảm xúc. Nó cũng có thể tự động học các biểu diễn cảm xúc từ dữ liệu lớn, từ đó cải thiện khả năng dự đoán và phân loại cảm xúc. Điều này có nghĩa là BERT không chỉ có thể phân tích cảm xúc dựa trên từ ngữ đơn giản, mà còn có thể hiểu được các yếu tố phức tạp hơn như ngữ cảnh và từ ngữ đa nghĩa. Điều này giúp BERT phân tích cảm xúc một cách chính xác hơn so với các mô hình truyền thống, đồng thời cung cấp một cái nhìn sâu hơn vào cảm xúc của người dùng. Ngoài ra, BERT còn có khả năng tự học hỏi từ dữ liệu. Điều này có nghĩa là nó có thể tự động học hỏi và cải thiện các biểu diễn cảm xúc từ dữ liệu lớn, giúp cải thiện khả năng dự đoán và phân loại cảm xúc. Từ đó, làm cho BERT trở thành một công cụ mạnh mẽ trong việc phân tích cảm xúc, giúp các nhà nghiên cứu và nhà phát triển có thể hiểu rõ hơn về cảm xúc của người dùng.

Phân tích cảm xúc đang trở thành một lĩnh vực nổi bật trong xử lý ngôn ngữ tự nhiên. BERT, một công cụ mạnh mẽ, đã được công nhận rộng rãi trong việc giải quyết thách thức này, dẫn đến sự phát triển mạnh mẽ của các biến thể BERT.

(Y. Zhang và cộng sự, 2020) đã tập trung vào việc tối ưu hóa BERT để cải thiện hiệu suất trong phân tích cảm xúc. Họ đã phát triển SBERT, một biến thể của BERT, nhằm mục đích tăng cường khả năng nhận biết và phân loại cảm xúc trong các bài viết ngắn trên mạng xã hội. Kết quả cho thấy SBERT vượt trội hơn so với BERT gốc và các phương pháp truyền thống khác. Ngoài ra, (Xin và cộng sự, 2021) và (T. Zhang và cộng sự, 2019) đã áp dụng BERT để phân tích cảm xúc trong các bình luận trên mạng xã hội và đã đạt được kết quả ấn tượng. So sánh với các mô hình truyền thống, BERT thể hiện độ chính xác và độ phức tạp cao hơn đáng kể. Không những thế, (Yinhan Liu và cộng sự, 2019) đã tiến hành thử nghiệm với BERT trong phân tích cảm xúc của các bình luận trên mạng xã hội. Kết quả đã chứng minh rằng BERT có khả năng nhận diện và phân loại cảm xúc với độ chính xác cao, vượt trội so với các mô hình truyền thống. Cuối cùng, (Qiao và cộng sự, 2022) đã so sánh hiệu suất của BiLSTM và BERT trong phân tích cảm xúc trong các văn bản. Kết quả cho thấy cả hai mô hình đều có hiệu suất tốt, nhưng BERT thường cho kết quả tốt hơn. Tất cả những nghiên cứu này đều chứng tỏ sự ưu việt của BERT trong việc phân tích cảm xúc, mở ra hướng đi mới cho lĩnh vực này.

Tổng quan, những kết quả từ các nghiên cứu liên quan đã chứng minh sức mạnh và tiềm năng của BERT trong việc phân tích cảm xúc. Sự phát triển và tinh chỉnh của BERT liên tục mở ra nhiều cơ hội mới cho nghiên cứu và ứng dụng trong tương lai, đồng thời định hình cách chúng ta tiếp cận và hiểu biết về cảm xúc của con người thông qua dữ liệu văn bản.

2.2.2. BiLSTM (Bidirectional Long Short-Term Memory)

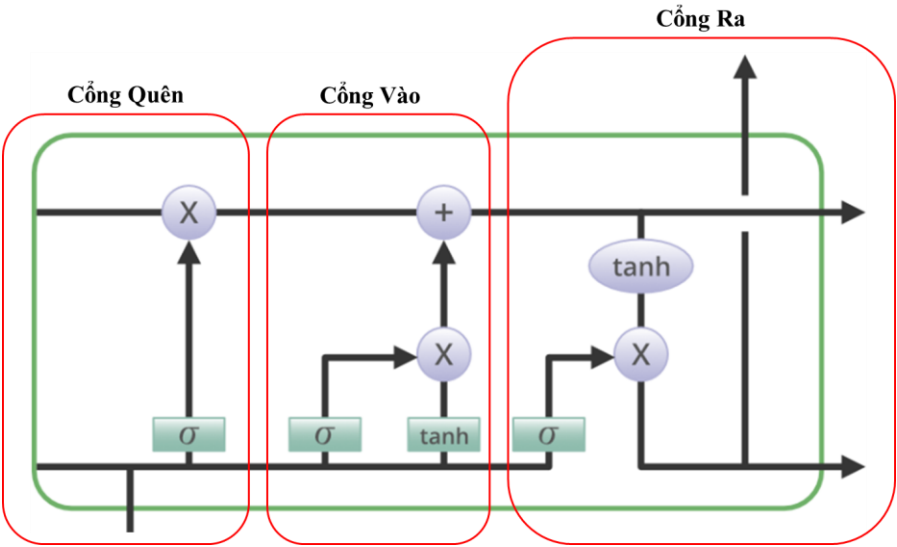
Mô hình BiLSTM (Bidirectional Long Short-Term Memory) là một biến thể của mạng LSTM (Long Short-Term Memory), một dạng mạng nơ-ron hồi quy sử dụng rộng rãi trong lĩnh vực xử lý ngôn ngữ tự nhiên và dữ liệu tuần tự. BiLSTM kết hợp hai lớp LSTM, mỗi lớp được thực hiện theo hai hướng khác nhau: một theo chiều thuận và một theo chiều ngược. Điều này cho phép mô hình "nhìn thấy" cả phần trước và phần sau của một từ trong câu hoặc một phần tử trong chuỗi dữ liệu.

Cấu trúc mạng BiLSTM

BiLSTM là một kiểu mạng nơ-ron sâu (deep neural network) dùng trong việc xử lý dữ liệu tuần tự như văn bản, âm thanh, hoặc dữ liệu dạng chuỗi khác. Đặc điểm nổi bật của BiLSTM là khả năng kết hợp thông tin từ cả hai phía của một từ hoặc một phần tử trong chuỗi dữ liệu. Cấu trúc của BiLSTM gồm hai phần chính:

- **LSTM thuận** (Forward LSTM): Đây là một chuỗi các mô-đun LSTM được sắp xếp theo thứ tự từ trái sang phải, nhận dữ liệu đầu vào theo thứ tự ban đầu của chuỗi.
- **LSTM ngược** (Backward LSTM): Tương tự như LSTM thuận, đây cũng là một chuỗi các mô-đun LSTM, nhưng được sắp xếp theo thứ tự ngược lại, từ phải sang trái, và nhận dữ liệu đầu vào theo chiều ngược lại của chuỗi.

Mỗi mô-đun LSTM trong BiLSTM bao gồm các cổng đầu vào (input gate), đầu ra (output gate), và cổng quên (forget gate). Các cổng này cho phép mạng học và điều chỉnh cách mà thông tin được truyền qua mạng, giúp mạng học được các mối quan hệ phức tạp trong dữ liệu. Được minh họa cụ thể qua **Hình 8**. Kết hợp giữa LSTM thuận và LSTM ngược cho phép BiLSTM học được thông tin từ cả hai hướng của một từ hoặc một phần tử trong chuỗi dữ liệu. Điều này giúp cải thiện khả năng hiểu biết và dự đoán trong các ứng dụng như phân tích cảm xúc, dịch máy, hoặc nhận diện đối tượng trong hình ảnh.



Hình 8. Cấu trúc Mạng LSTM.

Nhìn chung, nhờ vào khả năng hiểu rõ ngữ cảnh mà BiLSTM đóng vai trò quan trọng trong việc phân tích cảm xúc. Mô hình này có thể “nhìn” cả hai phía của một từ, giúp nắm bắt được ngữ cảnh xung quanh từ đó trong câu hoặc văn bản. Điều này giúp nâng cao hiệu suất của mô hình, cho phép dự đoán chính xác hơn cảm xúc được biểu đạt dựa trên ngữ cảnh rộng hơn. Ngoài ra, BiLSTM còn có khả năng xử lý dữ liệu tuần

tự, điều này rất hữu ích khi làm việc với dữ liệu ngôn ngữ tự nhiên, thường được biểu diễn dưới dạng chuỗi. Với khả năng “nhìn” cả hai phía của một từ, BiLSTM có thể mã hóa và phân tích cảm xúc từ các chuỗi dữ liệu như văn bản, đoạn hội thoại hoặc tweet một cách hiệu quả.

Một điểm khác cũng đáng chú ý là khả năng đánh giá toàn diện của BiLSTM. Bằng cách tính toán cả hai hướng, mô hình cung cấp thông tin từ cả phần trước và phía sau của một từ. Điều này giúp mô hình đánh giá toàn diện hơn về cảm xúc được biểu đạt trong một đoạn văn hay một đoạn hội thoại. Đồng thời, việc này cũng giúp mô hình có khả năng dự đoán chính xác hơn về cảm xúc trong các tình huống phức tạp và đa chiều.

Với khả năng đó của mình, mô hình BiLSTM đã trở thành một công cụ quan trọng trong lĩnh vực phân tích cảm xúc từ dữ liệu văn bản. Nhiều nghiên cứu đã khẳng định hiệu suất và tính hiệu quả của BiLSTM trong việc nhận diện và phân loại cảm xúc từ các đoạn văn ngắn trên mạng xã hội, các bình luận trực tuyến, hoặc đánh giá sản phẩm trên các sàn thương mại điện tử.

Đầu tiên, nghiên cứu của (Gaur và cộng sự, 2023) đã chỉ ra rằng mô hình BiLSTM có khả năng nhận diện và phân loại cảm xúc từ các đánh giá sản phẩm trên các sàn thương mại điện tử một cách chính xác và hiệu quả. Điều này mang lại lợi ích lớn cho các doanh nghiệp, giúp họ hiểu rõ ý kiến của khách hàng và điều chỉnh chiến lược kinh doanh. Tương tự, nghiên cứu của (Xu và cộng sự, 2019) đã khẳng định khả năng ưu việt của BiLSTM trong việc nhận diện và phân loại cảm xúc từ các bình luận trực tuyến. Điều này giúp tổ chức và các hệ thống trực tuyến dễ dàng phản hồi và xử lý phản hồi của người dùng một cách nhanh chóng và hiệu quả.

Tiếp đến, công trình của (Rhanoui và cộng sự, 2019) cũng tiếp tục củng cố vị thế của BiLSTM bằng cách so sánh hiệu suất của nó với các mô hình khác như CNN và SVM. Kết quả cho thấy sự ưu việt của BiLSTM trong việc phân loại cảm xúc, đặc biệt là trong các tác vụ đòi hỏi hiểu biết sâu sắc về ngôn ngữ tự nhiên và ngữ cảnh. Cuối cùng, nghiên cứu của (Hameed & Garcia-Zapirain, 2020) đã đề xuất một biến thể của BiLSTM được tinh chỉnh đặc biệt để phân tích cảm xúc trong các đoạn văn ngắn, và kết quả đã chỉ ra sự cải thiện đáng kể về độ chính xác trong dự đoán cảm xúc. Điều này thể hiện tiềm năng và sự linh hoạt của BiLSTM trong việc áp dụng vào các ứng dụng cụ thể và tăng cường hiệu suất của các hệ thống phân tích cảm xúc.

Những nghiên cứu liên quan đã chứng minh rằng mô hình BiLSTM đóng vai trò quan trọng và hiệu quả trong việc phân tích cảm xúc từ dữ liệu văn bản. Khả năng linh hoạt và tính toàn diện của mô hình đã mở ra nhiều cơ hội ứng dụng trong nhiều lĩnh vực, từ kinh doanh đến xã hội học và y học. Điều này làm tăng cường sự quan tâm và nghiên cứu sâu hơn về BiLSTM và các ứng dụng tiềm năng của nó trong thế giới thực.

2.2.3. RoBERTa (A Robustly Optimized BERT Pretraining Approach)

Các mô hình tự huấn luyện như BERT (Devlin và cộng sự, 2018) hay XLNet (Z. Yang và cộng sự, 2019) đã mang lại những thành tựu nổi bật trong lĩnh vực xử lý ngôn ngữ tự nhiên khi chúng đã nhanh chóng cải thiện được hiệu suất của nhiều tác vụ mà các mô hình học máy hay học sâu truyền thống phải mất rất nhiều tài nguyên để thực hiện. Tuy nhiên các nhà nghiên cứu của RoBERTa lại cho rằng BERT vẫn còn nhiều hạn

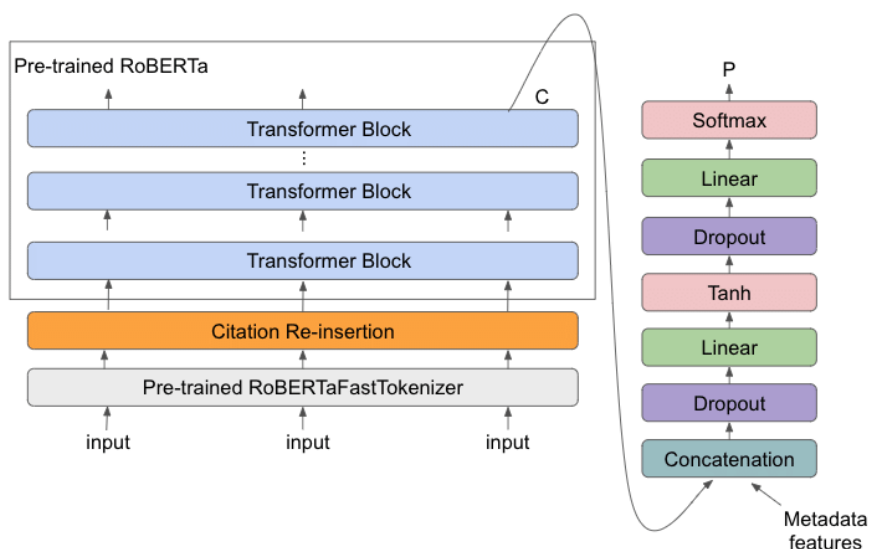
chế và chưa được huấn luyện một cách bài bản (Yinhan Liu và cộng sự, 2019). Vì thế nhóm tác giả đã đề xuất một phương pháp huấn luyện mới và RoBERTa được ra đời.

Về cơ bản, nhóm tác giả từ Facebook AI đã xây dựng lại dữ liệu huấn luyện cho mô hình BERT vì theo như quan sát của các nhà nghiên cứu thì hiệu suất của BERT sẽ cải thiện đáng kể nếu như được huấn luyện trên một tập dữ liệu lớn hơn. Tập dữ liệu này bao gồm kho tài liệu sau:

- BookCorpus + Wikipedia tiếng Anh (16GB) : Đây là dữ liệu mà BERT được đào tạo.
- CC-News (76GB) : Các tác giả đã thu thập dữ liệu này từ phần tiếng Anh của CommonCrawl News Data. Nó chứa 63 triệu tin bài tiếng Anh được thu thập từ tháng 9 năm 2016 đến tháng 2 năm 2019.
- OpenWebText (38GB) : Phiên bản mã nguồn mở của tập dữ liệu WebText được sử dụng để đào tạo OpenAI GPT .
- Câu chuyện (31GB) : Một tập hợp con dữ liệu CommonCrawl được lọc để phù hợp với phong cách giống câu chuyện của lược đồ Winograd.

Ngoài ra, một số cải tiến có thể kể đến như RoBERTa sử dụng chiến lược che lấp động, trong đó các token khác nhau được che lấp trong mỗi ví dụ huấn luyện. Điều này cho phép mô hình học được một tập hợp biểu diễn đa dạng hơn, vì nó phải dự đoán các mặt nạ khác nhau trong các ngữ cảnh khác nhau. RoBERTa cũng không sử dụng mask dự đoán câu tiếp theo (NSP) trong quá trình huấn luyện trước giống như BERT. Điều này cho phép RoBERTa tập trung hoàn toàn vào mục tiêu mô hình ngôn ngữ được che lấp, dẫn đến biểu diễn ngôn ngữ biểu cảm hơn.

Cấu trúc mạng RoBERTa



Hình 9. Cấu trúc của mô hình RoBERTa.

Giống như BERT, cấu trúc của RoBERTa được xây dựng dựa trên kiến trúc Transformer. Trong kiến trúc Transformer, cơ chế tự chú ý (self-attention) cho phép mô hình xác định mức độ quan trọng của các từ trong chuỗi đầu vào và tính toán biểu diễn sao cho có sự cân nhắc đến ngữ cảnh toàn bộ chuỗi. Các lớp feed-forward sau đó biến đổi các biểu diễn này thành biểu diễn cuối cùng của câu.

Cụ thể trong hình ảnh minh họa bên dưới cho thấy cấu trúc bao gồm các khối Transformer xếp chồng lên nhau. Mỗi khối Transformer xử lý chuỗi đầu vào và tạo ra chuỗi đầu ra. Mô hình được huấn luyện trước với một tập dữ liệu lớn đã được đề cập ở trên.

Khi xử lý văn bản đầu vào, lớp tokenizer sẽ chuyển đổi văn bản thành các token. Các khối Transformer xử lý các token này và tạo ra chuỗi token đầu ra. Lớp Positional Encoding thêm thông tin vị trí của các token vào mô hình. Lớp Dropout giúp mô hình tránh hiện tượng overfitting. Hàm Tanh phi tuyến hóa đầu ra của mô hình. Lớp Linear chuyển đổi đầu ra thành dạng mong muốn. Hàm Softmax tính toán xác suất của từng token trong chuỗi đầu ra.

Kết quả của một số nghiên cứu đã chứng minh được sự cải tiến đáng kể RoBERTa so với BERT dựa trên các thông số đánh giá cơ bản (Adoma và cộng sự, 2019). Ngoài ra, nghiên cứu của (Staliūnaitė & Iacobacci, 2020) cũng cho thấy sự cải thiện về mặt biểu diễn từ vựng tốt hơn của RoBERTa so với BERT tuy nhiên vấn đề ngữ nghĩa hợp thành không có sự khác biệt quá đáng kể của 2 mô hình.

2.3. Phân tích cảm xúc

Phân tích cảm xúc (Sentiment Analysis) là một phương pháp trong lĩnh vực xử lý ngôn ngữ tự nhiên nhằm mục đích phân loại văn bản theo các hình thái cảm xúc khác nhau. Khái niệm này lần đầu tiên được nhắc đến vào những năm thập niên 50, kỹ thuật này khi ấy được ứng dụng trong các tài liệu văn bản giấy truyền thống. Tuy nhiên, sự phát triển bùng nổ của thông tin dữ liệu số đã mở ra các ứng dụng hoàn toàn mới. Kỹ thuật phân tích cảm xúc giờ đây được áp dụng để tập trung nghiên cứu vào việc tính toán và phân loại ý kiến, thái độ và cảm xúc dưới nhiều hình thức khác nhau, bao gồm văn bản, hình ảnh và lời nói.

Nghiên cứu của (Kumar và cộng sự, 2023) đã đề cập có 3 mức độ chủ yếu của phân tích cảm xúc bao gồm:

- **Cấp độ tài liệu** (Document Level): Ở cấp độ này, toàn bộ tài liệu thể hiện ý kiến tích cực hoặc tiêu cực. Ví dụ, tài liệu đánh giá sản phẩm có thể chứa ý kiến tích cực hoặc tiêu cực về sản phẩm đó. Cấp độ này đại diện cho một ý kiến duy nhất trong tài liệu, do đó nó thuộc về cấp độ tài liệu.
- **Cấp độ câu** (Sentence Level): Đây là cấp độ thứ hai được sử dụng rộng rãi trên các trang web thương mại điện tử. Ở cấp độ này, mỗi câu được phân loại thành ý kiến tích cực, tiêu cực hoặc trung lập.
- **Cấp độ khía cạnh** (Aspect Level): Cấp độ này còn được gọi là phân tích dựa trên khía cạnh. Ở loại phân tích này, mỗi đánh giá được phân loại thành các khía cạnh và ý kiến liên quan đến chúng. Cấp độ này cung cấp thông tin chi tiết hơn về ý kiến tích cực hoặc tiêu cực đối với từng khía cạnh cụ thể. Ví dụ, câu nhận xét “Thức ăn rất ngon tại khách sạn” là một ví dụ về phân tích cảm xúc dựa trên khía cạnh, trong đó thức ăn là một khía cạnh của đánh giá.

Ngày nay, phân tích cảm xúc liên quan chặt chẽ đến lĩnh vực xử lý ngôn ngữ tự nhiên và trí tuệ nhân tạo. Do đó, đã xuất hiện nhiều phương pháp mới thay thế cho các phương pháp truyền thống, nhằm nâng cao hiệu quả và hiệu suất nghiên cứu. (Nandwani & Verma, 2021) đã đề cập đến ba kỹ thuật chính, bao gồm xử lý dựa trên từ vựng, mô hình học máy và mô hình học sâu và mỗi phương pháp đều có những ưu điểm và nhược điểm riêng. Tuy nhiên, trong bài báo này, nhóm tác giả sẽ chỉ tập trung đo lường tính toán hiệu quả hoạt động của các mô hình học sâu.

2.4. Ứng dụng

Phân tích cảm xúc đang ngày càng phát triển mạnh mẽ trong lĩnh vực nghiên cứu về xử lý ngôn ngữ tự nhiên, kéo theo đó là rất nhiều các ứng dụng thực tế trên các lĩnh vực khác nhau. Trong lĩnh vực nhà hàng và khách sạn, áp dụng phân tích cảm xúc có thể giúp khách hàng lựa chọn tốt hơn và cũng giúp chủ sở hữu cải thiện. Phân tích cảm xúc dựa trên các khía cạnh của khách sạn và nhà hàng sẽ giúp xác định khía cạnh có số lượng đánh giá tích cực và tiêu cực nhiều nhất, nhà cung cấp dịch vụ có thể đạt được lợi nhuận cao nhất khi họ có thể nhận biết được khía cạnh bị phản hồi tiêu cực nhiều nhất và cải thiện chúng tốt hơn (Akhtar và cộng sự, 2017).

Trong lĩnh vực chăm sóc sức khỏe, mạng xã hội trực tuyến như Twitter đã trở thành nguồn thông tin quan trọng liên quan đến sức khỏe. Công trình của (Clark và cộng sự, 2018) đã sử dụng các tweet liên quan đến trải nghiệm của bệnh nhân để phân tích sức khỏe cộng đồng. Trong vòng một năm, họ đã thu thập khoảng năm triệu tweet liên quan đến ung thư vú bằng cách sử dụng Twitter's Streaming API. Sau khi tiền xử lý, các tweet được phân loại bằng một mô hình phân loại hồi quy tuyến tính (Linear Regression) và một mô hình mạng nơ-ron tích chập (CNN). Các trải nghiệm tích cực về điều trị, sự hỗ trợ và việc tăng cường nhận thức của công chúng đều được liên kết. Tóm lại, việc áp dụng phân tích tình cảm để phân tích dữ liệu do bệnh nhân tạo ra trên mạng xã hội có thể giúp xác định nhu cầu và quan điểm của bệnh nhân.

Ngoài ra, phân tích cảm xúc còn đóng một vai trò quan trọng trong lĩnh vực dự đoán xu hướng biến động của thị trường cổ phiếu. Trong công trình của (Xing và cộng sự, 2018) cho thấy được xu hướng tăng giảm của cổ phiếu qua việc phân tích các thông tin mới nhất từ thị trường, dữ liệu được họ thu thập qua các nguồn khác nhau như Twitter, các bài báo, blogs, v.v và thực hiện phân tích cảm xúc ở mức câu đối với các văn bản này sẽ dự đoán được xu hướng của cổ phiếu trong tương lai. Cụ thể, giá cổ phiếu sẽ tăng khi xu hướng của thông tin thị trường trước đó mang chiều hướng tích cực và ngược lại.

3. Công trình nghiên cứu liên quan

Trong những năm gần đây, lĩnh vực phân tích cảm xúc dựa trên phương pháp học sâu đã thu hút sự quan tâm đặc biệt trong cộng đồng nghiên cứu khoa học. Điều này có nguyên nhân từ sự phát triển mạnh mẽ của mạng xã hội và dữ liệu trực tuyến, khiến việc hiểu và phân tích cảm xúc từ các nội dung trở nên càng quan trọng hơn bao giờ hết.

Một số nghiên cứu đáng chú ý đã tập trung vào việc áp dụng mô hình học sâu để xây dựng các hệ thống phân tích cảm xúc mạnh mẽ và hiệu quả. Ví dụ, công trình của (T. Chen và cộng sự, 2017) đã sử dụng mô hình học sâu để phân tích cảm xúc trong các bài viết trên mạng xã hội, và kết quả nghiên cứu này đã giúp hiểu rõ hơn về ngữ nghĩa, cấu trúc và cảm xúc trong các nội dung trên mạng xã hội.

Mô hình BERT (Bidirectional Encoder Representations from Transformers), được giới thiệu bởi (Cesar và cộng sự, 2023), cũng đã mở ra những tiềm năng mới trong xử lý ngôn ngữ tự nhiên và phân tích cảm xúc. Kết hợp với dữ liệu lớn từ internet, BERT đã cho thấy hiệu suất ưu việt trong việc phân tích cảm xúc. Các nghiên cứu như của (De Arriba và cộng sự, 2021) đã chứng minh hiệu suất xuất sắc của mô hình BERT trong việc nhận diện và phân loại cảm xúc trên mạng xã hội.

Ngoài ra, nhiều mô hình học sâu khác cũng đã được nghiên cứu và áp dụng trong phân tích cảm xúc trên mạng xã hội. Ví dụ, mô hình BiLSTM-CRF đề xuất bởi (Miao và cộng sự, 2021) đã thành công trong việc nhận diện cảm xúc trong các bình luận trực tuyến. Các mô hình biến thể khác như mạng nơ-ron của (Y. Chen, 2015), mô hình kết hợp chú ý tập trung của (Huang và cộng sự, 2021), và ứng dụng Capsule Network của (Dong và cộng sự, 2020) cũng đã được áp dụng trong phân tích cảm xúc trên mạng xã hội. Các phương pháp kết hợp học sâu và thống kê truyền thống cũng đã được áp dụng, như phương pháp của (L. Yang và cộng sự, 2020) trong việc phân tích cảm xúc từ đánh giá sản phẩm trên các sàn thương mại điện tử. Mô hình Recursive Neural Tensor Network (RNTN) của (Socher và cộng sự, 2013) cũng đã mở ra cánh cửa mới trong việc phân tích cảm xúc trong các câu văn phức tạp.

Đáng chú ý, (Gunawan và cộng sự, 2020) đã đề xuất một phương pháp kết hợp giữa mô hình học sâu và kỹ thuật tập trung nhằm cải thiện khả năng nhận diện cảm xúc và tập trung vào các phần quan trọng trong văn bản. Phương pháp này đã đem lại những đột phá đáng kể trong việc hiểu và phân loại cảm xúc từ các bài viết trên mạng xã hội. Bên cạnh việc phân tích cảm xúc từ các bài viết, nghiên cứu đã mở rộng phạm vi bằng cách áp dụng mô hình học sâu như LSTM (Long Short-Term Memory) và CNN (Convolutional Neural Network) để phân loại cảm xúc từ các bình luận và đánh giá trực tuyến. Để bổ sung, các nghiên cứu của (Hizlisoy và cộng sự, 2021) và tác giả (Cai & Xia, 2015) cũng đã chứng minh tính khả thi và độ chính xác của việc sử dụng mô hình học sâu trong việc phân loại cảm xúc từ dữ liệu trực tuyến, đặc biệt là từ hình ảnh và video.

Nhìn chung, việc áp dụng mô hình học sâu vào phân tích cảm xúc từ hình ảnh và video trên mạng xã hội cũng đã thu hút sự quan tâm của các nhà nghiên cứu. (Cai & Xia, 2015) đã thử nghiệm việc sử dụng mạng nơ-ron tích chập (CNN) để phân tích cảm xúc từ dữ liệu hình ảnh và video, mở ra tiềm năng mới trong việc nhận diện và phân loại cảm xúc đa phương tiện. (Mahendhiran & Kannimuthu, 2018) cũng đã đề xuất một phương pháp kết hợp học sâu từ văn bản và hình ảnh để phân tích cảm xúc trên mạng xã hội, mở ra những cơ hội để khai thác cảm xúc từ các nguồn dữ liệu đa phương tiện.

Tổng quan, các nghiên cứu trên đã chứng minh tính khả thi và hiệu quả của việc áp dụng mô hình học sâu và các kỹ thuật liên quan để phân tích cảm xúc từ dữ liệu trên mạng xã hội. Đồng thời, những nghiên cứu này đã mở ra những tiềm năng trong việc hiểu và khai thác cảm xúc từ các nguồn dữ liệu đa phương tiện.

Ứng dụng của học sâu trong phân tích cảm xúc không chỉ giới hạn trong lĩnh vực xã hội mạng và dữ liệu trực tuyến mà còn mở rộng sang nhiều lĩnh vực khác. Ví dụ, trong lĩnh vực thị trường tài chính, phân tích cảm xúc từ các bài viết, bình luận và thông điệp trên các diễn đàn tài chính có thể cung cấp thông tin quan trọng cho các nhà đầu tư về tình hình thị trường và xu hướng dự báo. Trong lĩnh vực chăm sóc sức khỏe, phân tích cảm xúc từ các đánh giá và bài viết của bệnh nhân có thể giúp xác định mức độ hài lòng và chất lượng dịch vụ y tế.

Tuy nhiên, việc phân tích cảm xúc dựa trên học sâu cũng đặt ra một số thách thức. Trước hết, việc xây dựng tập dữ liệu huấn luyện lớn và đa dạng là rất quan trọng để đảm bảo tính chính xác và tổng quát của mô hình. Bên cạnh đó, việc giải thích và hiểu rõ quyết định của mô hình học sâu cũng là một vấn đề đáng quan tâm. Mô hình học sâu thường được xem là "hộp đen" với khả năng học tự động mạnh mẽ, nhưng việc giải thích cụ thể làm thế nào và tại sao một quyết định được đưa ra có thể gây khó khăn.

Trong tương lai, lĩnh vực phân tích cảm xúc dựa trên học sâu có tiềm năng phát triển mạnh mẽ. Với sự tiến bộ của các mô hình học sâu, như mạng nơ-ron tái cấu trúc (RNN), mạng nơ-ron biến đổi (Transformer), và mạng nơ-ron tích chập (CNN), ta có thể tạo ra các mô hình phân tích cảm xúc mạnh mẽ hơn, hiệu quả hơn và linh hoạt hơn. Ngoài ra, việc kết hợp học sâu với các phương pháp khác như học tăng cường (Reinforcement Learning) và học máy không giám sát (Unsupervised Learning) có thể đem lại những tiến bộ đáng kể trong việc phân tích cảm xúc.

Tóm lại, lĩnh vực phân tích cảm xúc dựa trên học sâu đang phát triển mạnh mẽ và mang lại nhiều đóng góp quan trọng. Từ việc hiểu và phân tích cảm xúc trên mạng xã hội, đến áp dụng trong các lĩnh vực như tài chính và chăm sóc sức khỏe, học sâu đang mở ra những tiềm năng to lớn. Tuy vậy, còn nhiều thách thức phải đối mặt và cần có sự tiếp tục nghiên cứu và phát triển để khai thác hết tiềm năng của lĩnh vực này.

4. Nghiên cứu so sánh

4.1. Bộ dữ liệu

4.1.1. Zomato Bangalore Restaurants

Bộ dữ liệu **Zomato Bangalore Restaurants** được thu thập và công bố trên hai nền tảng chính là Kaggle và GitHub bởi Himanshu Poddar. Mục tiêu chính của bộ dữ liệu này là cung cấp thông tin chi tiết về các nhà hàng tại Bengaluru. Bộ dữ liệu này tập trung vào việc phân tích để giúp người dùng hiểu rõ hơn về các yếu tố ảnh hưởng đến quy trình vận hành và phản hồi của khách hàng đối với các nhà hàng trong khu vực này.

Bộ dữ liệu bao gồm tổng cộng 56,251 quan sát, trong đó mỗi quan sát đại diện cho một bình luận nhận xét về một nhà hàng cụ thể trong khu vực, đi kèm với các thông tin liên quan. Để mô tả rõ hơn, tác giả đã sử dụng tổng cộng 17 biến thuộc tính, tương ứng với các thông tin cụ thể về cả khách hàng và nhà hàng, nhằm cung cấp cái nhìn trực quan nhất về bộ dữ liệu.

Tên thuộc tính	Mô tả	Chú thích
url	Đường dẫn liên kết tới trang của nhà hàng	
address	Địa chỉ của nhà hàng	
name	Tên nhà hàng	
online_order	Phân loại nhà hàng chấp nhận sử dụng đặt hàng online	
book_table	Phân loại nhà hàng chấp nhận đặt bàn	
rate	Điểm đánh giá trung bình trên thang điểm 5	

votes	Số lượng khách hàng tham gia đánh giá	
phone	Số điện thoại nhà hàng nếu có	
location	Khu phố nơi đặt địa chỉ của nhà hàng	
Srest_type	Loại hình nhà hàng	
dish_liked	Những món ăn đề xuất được yêu thích tại nhà hàng	
cuisines	Loại ẩm thực được phục vụ tại nhà hàng	
approx_cost(for two people)	Chi phí xấp xỉ cho hai người	
reviews_list	Danh sách đánh giá về nhà hàng từ người dùng	Với mỗi đánh giá sẽ bao gồm hai phần: - Điểm xếp hạng; - Nội dung đánh giá từ người dùng.
menu_item	Danh sách món ăn trong thực đơn nhà hàng nếu có	
listed_in(type)	Loại bữa ăn	
listed_in(city)	Thành phố nơi nhà hàng được liệt kê	

4.1.2. Restaurant Reviews

Trong bài báo này, nhóm nghiên cứu đã sử dụng một bộ dữ liệu thứ hai mang tên “Restaurant Reviews”, được tác giả Arsh Anwar công bố trên nền tảng Kaggle. Bộ dữ liệu này là một bộ dữ liệu mô phỏng, bao gồm các đánh giá của khách hàng về các dịch vụ và món ăn tại nhà hàng, cùng với thông tin về việc họ có thích trải nghiệm đó hay không.

So với bộ dữ liệu đầu tiên, bộ dữ liệu mô phỏng này có số lượng quan sát ít hơn, chỉ tổng cộng 1,000 dòng và hai biến thuộc tính chính. Mỗi quan sát đại diện cho một nhận xét về nhà hàng. Nhờ các đánh giá khách quan và chi tiết, bộ dữ liệu này cung cấp một cơ sở dữ liệu lý tưởng cho các nghiên cứu về chất lượng dịch vụ nhà hàng và ẩm thực.

Qua việc phân tích bộ dữ liệu này, nhóm nghiên cứu có thể so sánh hiệu suất của các mô hình học sâu mà nhóm đã xây dựng với các bộ dữ liệu khác nhau trong cùng lĩnh vực. Đồng thời, nhóm nghiên cứu cũng có thể khai thác các ưu điểm và nhược điểm của từng mô hình bằng cách kiểm tra sự khác biệt giữa các thuộc tính của bộ câu sử dụng ngôn ngữ tự nhiên từ cả hai bộ dữ liệu này. Điều này giúp nhóm nghiên cứu hiểu rõ hơn về cách thức hoạt động và hiệu quả của các mô hình trong việc xử lý và phân tích dữ liệu.

Tên thuộc tính	Mô tả	Chú thích
Review	Chi tiết nội dung đánh giá từ người dùng	
Liked	Đánh giá cuối cùng của nhận xét.	Bao gồm hai giá trị chia đều cho tổng số quan sát: - “1”: tích cực; - “0”: tiêu cực.

4.2. Phương pháp luận

Nghiên cứu này tập trung vào việc khám phá và so sánh hiệu suất của các mô hình học sâu trong việc phân tích cảm xúc từ hai bộ dữ liệu khác nhau trong cùng một lĩnh vực. Mục tiêu chính là nắm bắt rõ hơn về các yếu tố ảnh hưởng đến quá trình đánh giá và phản hồi của người dùng đối với các nhà hàng, đồng thời so sánh và đánh giá hiệu suất của các mô hình học sâu trong việc dự đoán cảm xúc từ các đánh giá nhà hàng trực tuyến.

Quy trình nghiên cứu được thực hiện như sau:

- Thu thập và tiền xử lý dữ liệu:** Tiến hành thu thập và tiền xử lý dữ liệu từ hai bộ dữ liệu được chọn.
- Xây dựng và huấn luyện mô hình:** Xây dựng và huấn luyện các mô hình học sâu, bao gồm các mô hình như Bi-LSTM, BERT và RoBERTa.
- Đánh giá mô hình:** Mỗi mô hình được đánh giá dựa trên các tiêu chí như độ chính xác, độ phân loại, và các độ đo khác nhau để đánh giá hiệu suất. Điều này giúp đo lường khả năng dự đoán của mô hình và đưa ra cái nhìn toàn diện về hiệu suất của từng mô hình.
- Thực hiện thử nghiệm:** Thực hiện một loạt các thử nghiệm để so sánh hiệu suất của các mô hình trên hai bộ dữ liệu khác nhau và đánh giá sự khác biệt trong kết quả. Điều này giúp hiểu rõ hơn về khả năng tổng quát hóa của các mô hình trên dữ liệu từ các nguồn khác nhau và đưa ra những nhận định về sự thích ứng và linh hoạt của chúng.
- Phân tích kết quả và rút ra kết luận:** Sau khi thu được kết quả từ các phương pháp đánh giá và thử nghiệm, tiến hành phân tích kết quả và rút ra những kết luận về sự hiệu quả của các mô hình học sâu trong việc phân tích cảm xúc từ dữ liệu nhà hàng trực tuyến. Điều này giúp cung cấp cái nhìn sâu sắc hơn về khả năng ứng dụng của các phương pháp học sâu trong việc hiểu và phân tích ý kiến và cảm xúc từ dữ liệu ngôn ngữ tự nhiên trên mạng.

4.3. Phân loại cảm xúc

Trong phần này, phân tích về quá trình phân loại cảm xúc sử dụng các mô hình học sâu được thực hiện. Đầu tiên, tiến hành tiền xử lý dữ liệu bằng cách loại bỏ các thành phần không liên quan và thực hiện chuẩn hóa câu, giúp cải thiện hiệu suất huấn luyện mô hình. Các bước tiền xử lý các văn bản đánh giá thực hiện bao gồm các bước sau:

1. **Định dạng dữ liệu phân tích:** Thực hiện tạo dataframe mới chỉ chứa các dòng đánh giá (review) và nhãn phân loại (label). Đối với bộ dữ liệu chưa có nhãn phân loại cho nội dung đánh giá, sẽ sử dụng điểm đánh giá (rating) để định nghĩa các nhãn cho nội dung này.
2. **Loại bỏ đánh giá trùng lặp:** Các đánh giá trùng lặp được loại bỏ dựa trên nội dung của chúng, với các tiêu chí như độ dài tối thiểu và số lượng xuất hiện tối thiểu.
3. Chuyển đổi văn bản thành chữ thường.
4. Loại bỏ các thẻ HTML bằng cách sử dụng biểu thức chính quy để tìm và thay thế các chuỗi bắt đầu bằng '<' và kết thúc bằng '>'.
5. Loại bỏ các ký tự không phải là chữ cái hoặc số, thay thế chúng bằng khoảng trắng.
6. Tách văn bản thành các từ riêng biệt và loại bỏ từ dừng (stop words) tiếng Anh.
7. Stemming các từ về dạng gốc.
8. Thay thế tất cả các số trong văn bản bằng từ khóa 'NUM'.
9. Loại bỏ các hàng chứa dữ liệu không hợp lệ sau quá trình xử lý văn bản.

Ngoài ra, sau khi hoàn thành quá trình tiền xử lý dữ liệu văn bản, cũng thực hiện tiến hành một bước quan trọng khác là cân bằng mẫu dữ liệu. Điều này được thực hiện để giảm thiểu hiện tượng mất cân bằng trong số lượng các lớp của biến mục tiêu, từ đó tạo điều kiện thuận lợi cho việc huấn luyện mô hình dự đoán. Trong phương pháp này, sử dụng kỹ thuật tái chọn mẫu ngẫu nhiên (Random Under Sampling) từ thư viện imbalanced-learn để giảm bớt số lượng mẫu của lớp thiểu số để phù hợp với số lượng mẫu của lớp đa số.

Để thực hiện, trích xuất các đặc trưng (feature) và biến mục tiêu (target) từ dữ liệu ban đầu. Tiếp theo, sử dụng hàm RandomUnderSampler để thực hiện tái chọn mẫu ngẫu nhiên, với tùy chọn tái chọn mẫu `sampling_strategy = 1`. Sau khi tái chọn, thực hiện tạo lại cấu trúc dữ liệu bằng cách tạo một DataFrame mới từ dữ liệu đã tái chọn mẫu, với các cột được đặt tên tương ứng với đặc trưng và biến mục tiêu. Điều này giúp tập dữ liệu được cân bằng và sẵn sàng cho quá trình huấn luyện mô hình.

Sau khi bộ dữ liệu đã sẵn sàng, thực hiện quá trình phân loại cảm xúc sử dụng các mô hình học sâu. Mỗi mô hình được huấn luyện trên bộ dữ liệu lúc này đã qua xử lý và cân bằng mẫu. Việc này giúp tăng khả năng học của mô hình và cải thiện độ chính xác trong việc dự đoán cảm xúc của đánh giá.

Sau khi huấn luyện xong, đánh giá hiệu suất của mỗi mô hình dựa trên các tiêu chí như độ chính xác (accuracy), độ phân loại (classification rate) và các độ đo khác. Từ đó đánh giá khả năng dự đoán của từng mô hình và chọn ra mô hình có hiệu suất tốt nhất cho việc phân loại cảm xúc từ đánh giá nhà hàng trực tuyến. Ngoài ra, cũng thực hiện các thử nghiệm để so sánh hiệu suất của các mô hình trên hai bộ dữ liệu khác nhau và đánh giá sự khác biệt trong kết quả.

Cuối cùng, sau khi thu được kết quả từ các phương pháp đánh giá và thử nghiệm, tiến hành phân tích kết quả và rút ra những kết luận về sự hiệu quả của các mô hình học sâu trong việc phân tích cảm xúc, mục tiêu mong muốn sẽ được cung cấp cái nhìn sâu sắc hơn về khả năng ứng dụng của các phương pháp học sâu trong việc hiểu và phân tích ý kiến và cảm xúc từ dữ liệu ngôn ngữ tự nhiên trên mạng.

4.4. Xây dựng mô hình phân loại cảm xúc

4.4.1. Mô hình BERT

Mô hình BertForSequenceClassification được xây dựng dựa trên kiến trúc mạng nơ-ron BertModel, một biến thể của mô hình Transformer. Mô hình này bao gồm các phần chính như sau:

- Phần Embeddings chứa các lớp nhúng (embeddings) cho từng thành phần của dữ liệu đầu vào. Đầu tiên, một lớp nhúng từ được sử dụng để biểu diễn từ vựng, với kích thước ma trận là 30522×768 . Tiếp theo, các lớp nhúng vị trí và loại token cũng được sử dụng để bổ sung thông tin về vị trí của từ trong câu và loại của từ. Cuối cùng, một lớp dropout được áp dụng để tránh overfitting.
- Phần Encoder chứa một chuỗi các lớp BertLayer, mỗi lớp này bao gồm một phần Attention, Intermediate và Output. Trong phần Attention, một cơ chế tự chú ý (self-attention mechanism) được sử dụng để tạo ra các biểu diễn đặc trưng cho từng từ dựa trên các từ khác trong câu. Phần Intermediate và Output là các lớp mạng nơ-ron hoạt động để học các biểu diễn cấp cao hơn của dữ liệu.
- Phần Pooler chứa một lớp BertPooler được sử dụng để tạo ra biểu diễn tổng thể của cả câu dựa trên các biểu diễn từ được tạo ra từ phần Encoder. Lớp này giúp mô hình tìm ra biểu diễn tổng thể của câu để sử dụng cho việc phân loại câu.
- Phần Classifier bao gồm một lớp kết nối đầy đủ với hàm kích hoạt softmax, được sử dụng để phân loại dữ liệu thành hai nhãn lớp. Mô hình cuối cùng được huấn luyện bằng cách sử dụng dữ liệu đã được mã hóa và các tham số huấn luyện được cài đặt trong các đối số huấn luyện (TrainingArguments). Sau đó, một bộ Trainer được tạo để thực hiện quá trình huấn luyện với mô hình và các đối số huấn luyện đã được cung cấp.

4.4.2. Mô hình RoBERTa

Giống như BERT, mô hình Roberta được xây dựng dựa trên kiến trúc mạng nơ-ron RobertaModel, cũng là một biến thể của mô hình Transformer. Mô hình này bao gồm các phần chính như Embeddings, Encoder và Classifier.

Phần Embeddings chứa các lớp nhúng cho từng thành phần của dữ liệu đầu vào. Đầu tiên, một lớp nhúng từ được sử dụng để biểu diễn từ vựng, với kích thước ma trận là 50265×768 . Tiếp theo, các lớp nhúng vị trí và loại token cũng được sử dụng để bổ sung thông tin về vị trí của từ trong câu và loại của từ. Cuối cùng, một lớp dropout được áp dụng để tránh overfitting.

Phần Encoder chứa một chuỗi các lớp RobertaLayer, mỗi lớp này bao gồm một phần Attention, Intermediate và Output tương tự như mô hình BERT.

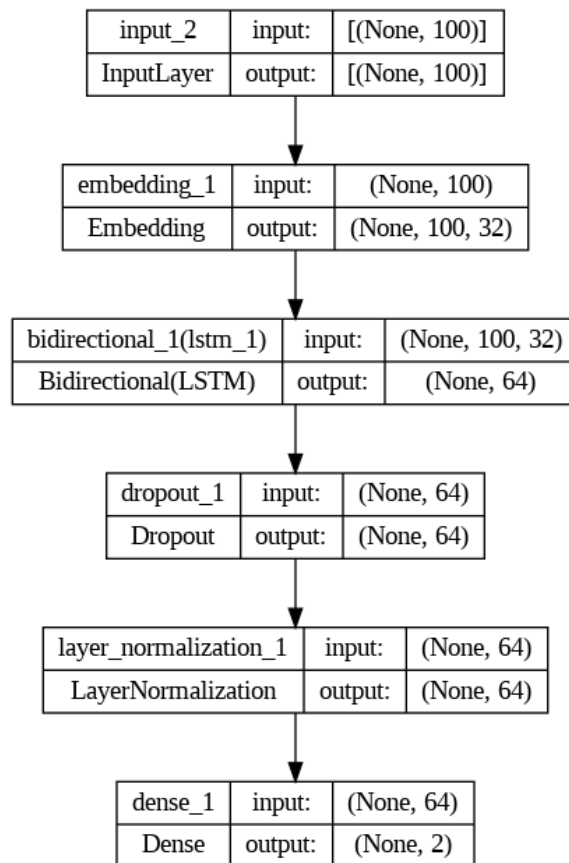
Cuối cùng là phần Classifier bao gồm một lớp kết nối đầy đủ với hàm kích hoạt softmax, được sử dụng để phân loại dữ liệu thành hai lớp ("positive" và "negative"). Mô hình cuối cùng được huấn luyện bằng cách sử dụng dữ liệu đã được mã hóa và các tham số huấn luyện được cài đặt trong các đối số huấn luyện (TrainingArguments).

4.4.3. Mô hình Bi-LSTM

Mô hình Bi-LSTM được xây dựng sử dụng kiến trúc mạng nơ-ron Long Short-Term Memory (LSTM) đa chiều. Dữ liệu được xử lý và chuẩn bị bằng cách sử dụng Tokenizer để mã hóa các văn bản thành chuỗi số và sử dụng phép đệm để đảm bảo rằng các chuỗi đầu vào có cùng độ dài, với giá trị maxlen_vector được thiết lập là 100.

Khởi đầu, mô hình bắt đầu bằng một lớp Embedding, các từ được biểu diễn dưới dạng vector với output dim là 32. Sau đó, một lớp Bi-LSTM được áp dụng với mức độ dropout là 0.2, nhằm học các biểu diễn đặc trưng của dữ liệu. Các lớp dropout này cũng áp dụng để tránh overfitting và sau đó sẽ có một lớp chuẩn hóa Normalization để cân bằng đầu ra. Cuối cùng sẽ có một lớp Dense với hàm kích hoạt softmax để phân loại dữ liệu thành hai lớp (“positive” và “negative”).

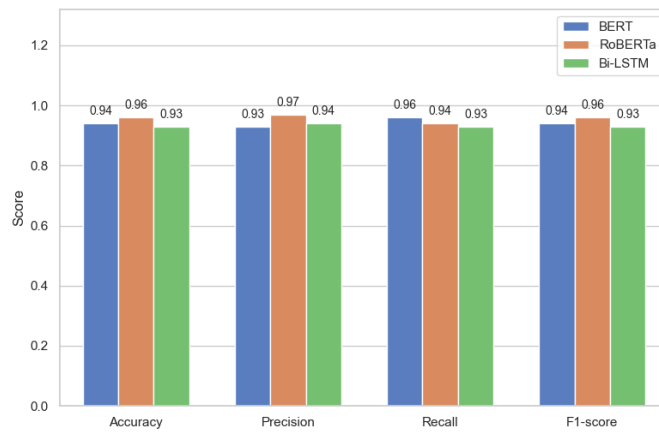
Mô hình được biên dịch với trình tối ưu hóa Adam và hàm mất mát nhị phân cross-entropy, chỉ định epochs = 20 để huấn luyện mô hình, với kích thước batch là 128 và tập dữ liệu được chia thành phần validation với tỷ lệ 10% để đánh giá hiệu suất. Mô hình sau cùng được lưu lại qua quá trình huấn luyện bằng cách sử dụng hàm callback ModelCheckpoint.



Hình 10. Chi tiết kiến trúc của mô hình Bi-LSTM áp dụng trong bài nghiên cứu.

5. Kết quả thực nghiệm

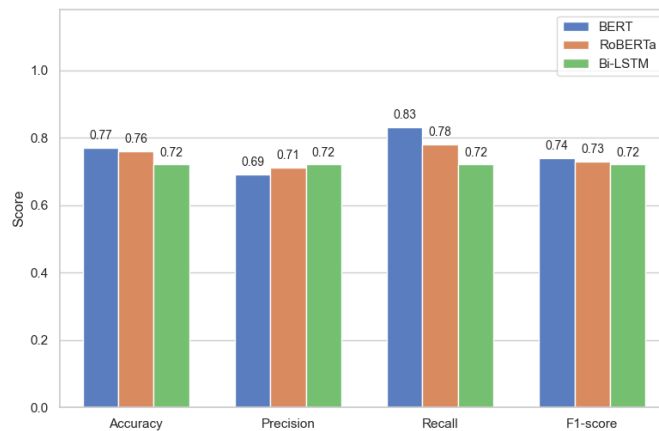
Trong nghiên cứu này, các thử nghiệm đã được tiến hành để đánh giá hiệu suất của các mô hình đã nêu trên trên các bộ dữ liệu khác nhau. Sau khi đánh giá tất cả ba mô hình dựa trên các chỉ số đánh giá, các biểu đồ đã được tạo ra để thể hiện kết quả một cách trực quan, giúp việc so sánh trở nên dễ dàng hơn.



Hình 11. Biểu đồ cột ghép thể hiện các chỉ số đánh giá từng mô hình của bộ dữ liệu Zomato Bangalore Restaurants

Đã nhận thấy rằng cả hai mô hình BERT và RoBERTa đều cho thấy hiệu suất cao, với độ chính xác lần lượt là 0.94 và 0.96, chỉ có sự chênh lệch nhỏ giữa chúng. Tuy nhiên, khi xem xét các chỉ số khác như Precision, Recall và F1-score, đã nhận thấy mô hình RoBERTa có vẻ hiệu quả hơn và cho thấy sự cân bằng giữa Precision và Recall tốt hơn so với BERT. Trong khi đó, một mô hình học sâu khác là Bi-LSTM cũng có độ chính xác khá cao và chỉ thấp hơn nhóm mô hình transformer này không đáng kể (0.93).

Tương tự đối với bộ dữ liệu thứ hai, cũng đã tạo ra biểu đồ cột để đánh giá các chỉ số đánh giá của ba mô hình.



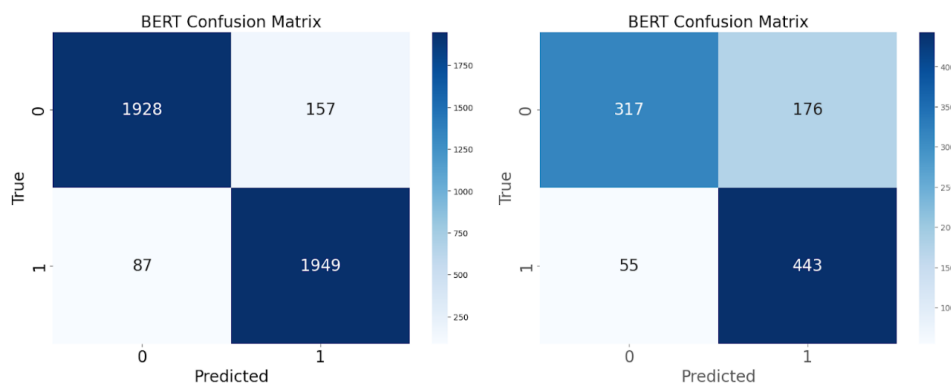
Hình 12. Biểu đồ cột ghép thể hiện các chỉ số đánh giá từng mô hình của bộ dữ liệu Restaurant Reviews

Thông qua các biểu đồ trực quan, cả hai mô hình BERT và RoBERTa đều cho thấy độ chính xác cao, đạt lần lượt là 0.77 và 0.76, chỉ chênh lệch nhỏ giữa chúng. Tuy nhiên, khi xem xét Precision, Recall và F1-score, RoBERTa có xu hướng hiển thị hiệu suất tốt hơn so với BERT. Cụ thể, RoBERTa có Precision, Recall và F1-score lần lượt là 0.71, 0.78 và 0.73, trong khi BERT có các giá trị tương ứng là 0.69, 0.83 và 0.74. Điều này cho thấy RoBERTa có sự cân bằng hơn giữa Precision và Recall so với BERT. Trong khi đó, mô hình Bi-LSTM có độ chính xác thấp hơn (khoảng 0.72) so với cả BERT và RoBERTa, nhưng các giá trị Precision, Recall và F1-score của nó đều ổn định ở mức 0.72.

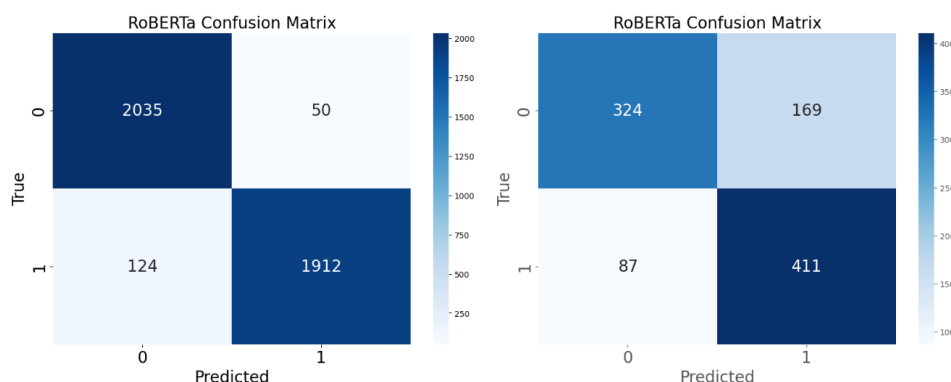
Trung bình, các chỉ số này đều ở mức tương đối cao, với tất cả các giá trị đều trên 70%. Tuy nhiên, tất cả đều thấp hơn nhiều so với bộ dữ liệu trước đó. Điều này có thể được giải thích bằng việc các mô hình đã

được huấn luyện trên bộ dữ liệu trước đó và sau đó được đánh giá trên bộ dữ liệu mới. Sự khác biệt giữa các bộ dữ liệu có thể dẫn đến hiệu suất giảm do các siêu tham số không tối ưu cho bộ dữ liệu mới.

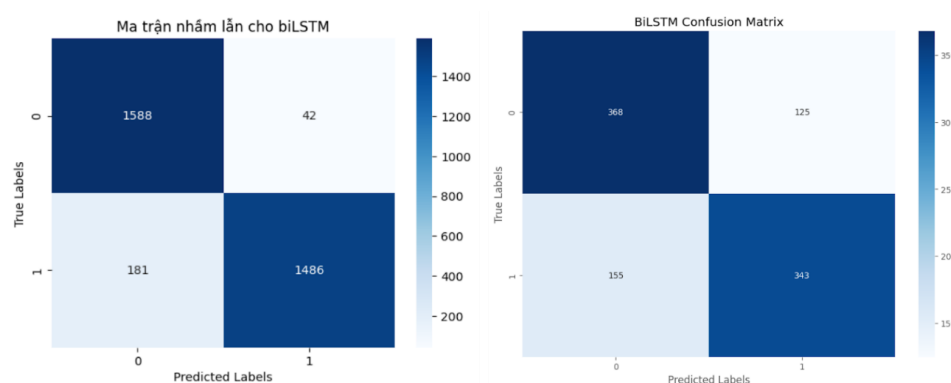
Để có cái nhìn rõ ràng hơn, đã thực hiện trực quan hóa thông qua ma trận nhầm lẫn để đưa ra nhận định chi tiết hơn.



Hình 13. Kết quả ma trận nhầm lẫn mô hình BERT của bộ dữ liệu Zomato Bangalore Restaurants (trái) và Restaurant Reviews (phải).



Hình 14. Kết quả ma trận nhầm lẫn mô hình RoBERTa của bộ dữ liệu Zomato Bangalore Restaurants (trái) và Restaurant Reviews (phải).



Hình 15. Kết quả ma trận nhầm lẫn mô hình Bi-LSTM của bộ dữ liệu Zomato Bangalore Restaurants (trái) và Restaurant Reviews (phải).

Dựa trên ma trận nhầm lẫn của ba mô hình BERT, RoBERTa và Bi-LSTM cho cả hai bộ dữ liệu, có thể quan sát được rằng phân loại của các mô hình có xu hướng tương đối tốt, nhưng có sự thay đổi đáng kể trong việc dự đoán nhầm lẫn giữa hai bộ dữ liệu.

Đối với mô hình BERT, đã nhận thấy rằng trong bộ dữ liệu thứ nhất (Zomato Bangalore Restaurants), có tổng cộng 244 dự đoán sai (157 dự đoán thuộc sai lầm loại I và 87 thuộc sai lầm loại II), chiếm 5.92% tổng số quan sát; trong khi trong bộ dữ liệu thứ hai (Restaurant Reviews), tổng số dự đoán sai tăng lên thành 251 (176 dự đoán thuộc sai lầm loại I và 55 thuộc sai lầm loại II), chiếm 24.83% số quan sát. Điều này cho thấy một sự tăng đáng kể trong việc dự đoán sai khi sử dụng bộ dữ liệu thứ hai.

Tương tự, với mô hình RoBERTa, trong bộ dữ liệu đầu tiên, có 4.22% trong tổng số quan sát bị dự đoán sai, 50 dự đoán cho sai lầm loại I và 124 dự đoán cho sai lầm loại II, trong khi trong bộ dữ liệu thứ hai, tổng số dự đoán sai tăng lên thành 25.83%, lần lượt cho hai sai lầm là 169 và 87 dự đoán.

Cuối cùng là mô hình Bi-LSTM. Trong bộ dữ liệu thứ nhất, có tổng cộng 223 dự đoán sai (42 dự đoán negative thành positive và 181 dự đoán positive thành negative), chiếm khoảng 7.05% tổng quan sát; trong khi ở bộ thứ hai, tổng số dự đoán sai tăng lên thành 280 (125 dự đoán negative thành positive và 155 dự đoán positive thành negative), chiếm khoảng 28.25%, cao nhất trong cả ba mô hình.

Qua thống kê phân tích, đã phát hiện ra một số điểm quan trọng cần nhận xét từ phân tích của ma trận nhầm lẫn cho ba mô hình BERT, RoBERTa và Bi-LSTM đối với cả hai bộ dữ liệu:

Mô hình BERT thường có xu hướng dự đoán sai lầm loại I nhiều hơn là sai lầm loại II cho cả hai bộ dữ liệu. Điều này có thể cho thấy rằng mô hình có khả năng dự đoán nhầm các trường hợp thuộc lớp negative thành positive nhiều hơn là dự đoán nhầm các trường hợp thuộc lớp positive thành negative. Ngược lại, cả hai mô hình RoBERTa và Bi-LSTM đều có xu hướng dự đoán sai lầm loại II nhiều hơn là sai lầm loại I. Điều này có thể ám chỉ rằng chúng có xu hướng dự đoán nhầm các trường hợp thuộc lớp positive thành negative nhiều hơn là dự đoán nhầm các trường hợp thuộc lớp negative thành positive.

Ngoài ra, mô hình RoBERTa cho thấy hiệu suất tốt nhất, với tỷ lệ dự đoán sai thấp nhất và gần nhất với mô hình BERT. Sự chênh lệch trong hiệu suất giữa RoBERTa và BERT không quá lớn, cho thấy cả hai mô hình có khả năng phân loại tương đương nhau.

Cuối cùng, mô hình Bi-LSTM có tỷ lệ dự đoán sai lầm cao nhất trong cả ba mô hình, đặc biệt là trong bộ dữ liệu thứ hai. Sự chênh lệch đáng kể này có thể cho thấy rằng mô hình Bi-LSTM không đạt được hiệu suất tốt nhất cho các bộ dữ liệu này và cần được cải thiện hoặc điều chỉnh để đảm bảo kết quả tốt hơn.

Dựa trên những phân tích và kết quả thu được, có thể rút ra một số kết luận quan trọng. Cả ba mô hình BERT, RoBERTa và Bi-LSTM đều cho thấy hiệu suất tốt trên cả hai bộ dữ liệu. Tuy nhiên, mô hình RoBERTa cho thấy hiệu suất tốt nhất, với tỷ lệ dự đoán sai thấp nhất và gần nhất với mô hình BERT. Trong khi đó, mô hình Bi-LSTM có tỷ lệ dự đoán sai lầm cao nhất trong cả ba mô hình, đặc biệt là trong bộ dữ liệu thứ hai. Điều này cho thấy rằng mô hình Bi-LSTM có thể cần được cải thiện hoặc điều chỉnh để đảm bảo kết quả tốt hơn. Dự định trong tương lai, có thể tập trung vào việc khám phá các phương pháp kết hợp (hybrid learning), nhằm nâng cao độ chính xác của phân loại cảm xúc và giảm chi phí tính toán. Mục tiêu là mở rộng phạm vi nghiên cứu bằng cách bao gồm cả các phương pháp và dữ liệu mới, cũng như đánh giá độ tin cậy và thời gian xử lý của các mô hình kết hợp trên nhiều loại dữ liệu khác nhau.

6. Kết luận

Trong nghiên cứu này, đã tiến hành các thử nghiệm để đánh giá hiệu suất của các mô hình BERT, RoBERTa và Bi-LSTM trên hai bộ dữ liệu khác nhau là Zomato Bangalore Restaurants và Restaurant Reviews.

Đối với bộ dữ liệu Zomato Bangalore Restaurants, cả hai mô hình BERT và RoBERTa đều cho thấy hiệu suất cao. Tuy nhiên, RoBERTa có vẻ hiệu quả hơn với sự cân bằng giữa Precision và Recall tốt hơn so với BERT. Mô hình Bi-LSTM cũng có độ chính xác khá cao. Còn trên bộ dữ liệu Restaurant Reviews, cả hai mô hình BERT và RoBERTa cũng cho thấy độ chính xác cao. Tương tự như trên bộ dữ liệu trước, RoBERTa hiển thị sự cân bằng giữa Precision và Recall tốt hơn so với BERT. Mô hình Bi-LSTM có độ chính xác thấp hơn một chút, nhưng vẫn ổn định với các giá trị Precision, Recall và F1-score đều ở mức ổn định.

Vậy từ nghiên cứu này, có thể kết luận rằng cả hai mô hình BERT và RoBERTa đã thể hiện hiệu suất cao trên cả hai bộ dữ liệu trong lĩnh vực F&B. Chi tiết hơn, RoBERTa có vẻ hiệu quả hơn khi so với BERT, trong khi mô hình Bi-LSTM cũng cho thấy độ chính xác khá cao và ổn định trên cả hai bộ dữ liệu nhưng không hiệu quả bằng hai mô hình transformer. Dự định trong các nghiên cứu tiếp theo, có thể tập trung vào việc khám phá các phương pháp kết hợp (hybrid learning), nhằm nâng cao độ chính xác của phân loại cảm xúc và giảm chi phí tính toán. Mục tiêu là mở rộng phạm vi nghiên cứu bằng cách bao gồm cả các phương pháp và dữ liệu mới, cũng như đánh giá độ tin cậy và thời gian xử lý của các mô hình kết hợp trên nhiều loại dữ liệu khác nhau.

Tài liệu tham khảo

- Acheampong, F. A., Nunoo-Mensah, H., & Chen, W. (2021). Transformer models for text-based emotion detection: a review of BERT-based approaches. *Artificial Intelligence Review*, 54(8), 5789-5829.
- Adoma, A. F., Henry, N.-M., & Chen, W. (2019, 2020). *Comparative analyses of bert, roberta, distilbert, and xlnet for text-based emotion recognition*.
- Akhtar, N., Zubair, N., Kumar, A., & Ahmad, T. (2017). Aspect based Sentiment Oriented Summarization of Hotel Reviews. *Procedia Computer Science*, 115, 563-571. doi:10.1016/j.procs.2017.09.115
- Batra, H., Pun, N. S., Sonbhadra, S. K., & Agarwal, S. (2021). *Bert-based sentiment analysis: A software engineering perspective*. Paper presented at the Database and Expert Systems Applications: 32nd International Conference, DEXA 2021, Virtual Event, September 27–30, 2021, Proceedings, Part I 32.
- Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: analyzing text with the natural language toolkit*. " O'Reilly Media, Inc."
- Cai, G., & Xia, B. (2015). *Convolutional neural networks for multimedia sentiment analysis*. Paper presented at the Natural Language Processing and Chinese Computing: 4th CCF Conference, NLPCC 2015, Nanchang, China, October 9-13, 2015, Proceedings 4.
- Cesar, L. B., Manso-Callejo, M.-Á., & Cira, C.-I. (2023). BERT (Bidirectional Encoder Representations from Transformers) for missing data imputation in solar irradiance time series. *Engineering Proceedings*, 39(1), 26.

- Chamekh, A., Mahfoudh, M., & Forestier, G. (2022). *Sentiment analysis based on deep learning in e-commerce*. Paper presented at the International Conference on Knowledge Science, Engineering and Management.
- Chen, D., & Manning, C. D. (2014). *A fast and accurate dependency parser using neural networks*. Paper presented at the Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP).
- Chen, T., Xu, R., He, Y., & Wang, X. (2017). Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN. *Expert Systems with Applications*, 72, 221-230.
- Chen, Y. (2015). *Convolutional neural network for sentence classification*. University of Waterloo,
- Clark, E., James, T., Jones, C., Alapati, A., Ukandu, P., Danforth, C., & Dodds, P. (2018). *A Sentiment Analysis of Breast Cancer Treatment Experiences and Healthcare Perceptions Across Twitter*.
- Dang, N. C., Moreno-García, M. N., & De la Prieta, F. (2020). Sentiment analysis based on deep learning: A comparative study. *Electronics*, 9(3), 483.
- De Arriba, A., Oriol, M., & Franch, X. (2021). *Applying transfer learning to sentiment analysis in social media*. Paper presented at the 2021 IEEE 29th International Requirements Engineering Conference Workshops (REW).
- Delgadillo, J., Kinyua, J., & Mutigwe, C. (2023). A BERT-Based Model for Financial Social Media Sentiment Analysis. *International Journal of Cognitive and Language Sciences*, 17(2), 112-117.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Dong, Y., Fu, Y., Wang, L., Chen, Y., Dong, Y., & Li, J. (2020). A sentiment analysis method of capsule network based on BiLSTM. *IEEE access*, 8, 37014-37020.
- Gaur, C., Kumar, A., & Das, S. (2023). *Analysis and Design of Document Similarity Using BiLSTM and BERT*. Paper presented at the International Conference on Advanced Communication and Intelligent Systems.
- Gunawan, T. S., Ashraf, A., Riza, B. S., Haryanto, E. V., Rosnelly, R., Kartiwi, M., & Janin, Z. (2020). Development of video-based emotion recognition using deep learning with Google Colab. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 18(5), 2463-2471.
- Hameed, Z., & Garcia-Zapirain, B. (2020). Sentiment classification using a single-layered BiLSTM model. *Ieee Access*, 8, 73992-74001.
- Hizlisoy, S., Yildirim, S., & Tufekci, Z. (2021). Music emotion recognition using convolutional long short term memory deep neural networks. *Engineering Science and Technology, an International Journal*, 24(3), 760-767.
- Huang, F., Li, X., Yuan, C., Zhang, S., Zhang, J., & Qiao, S. (2021). Attention-emotion-enhanced convolutional LSTM for sentiment analysis. *IEEE transactions on neural networks and learning systems*, 33(9), 4332-4345.

- Kumar, S., Roy, P. P., Dogra, D. P., & Kim, B.-G. (2023). A Comprehensive Review on Sentiment Analysis: Tasks, Approaches and Applications. *arXiv preprint arXiv:2311.11250*.
- Lample, G., Ballesteros, M., Subramanian, S., Kawakami, K., & Dyer, C. (2016). Neural architectures for named entity recognition. *arXiv preprint arXiv:1603.01360*.
- Li, X., Lei, Y., & Ji, S. (2022). BERT-and BiLSTM-based sentiment analysis of online Chinese buzzwords. *Future Internet*, 14(11), 332.
- Liu, B., & Zhang, L. (2012). A survey of opinion mining and sentiment analysis. In *Mining text data* (pp. 415-463): Springer.
- Liu, Y., Lu, J., Yang, J., & Mao, F. (2020). Sentiment analysis for e-commerce product reviews by deep learning model of Bert-BiGRU-Softmax. *Mathematical Biosciences and Engineering*, 17(6), 7819-7837.
- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., . . . Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.
- Lovins, J. B. (1968). Development of a stemming algorithm. *Mech. Transl. Comput. Linguistics*, 11(1-2), 22-31.
- Mahendhiran, P., & Kannimuthu, S. (2018). Deep learning techniques for polarity classification in multimodal sentiment analysis. *International Journal of Information Technology & Decision Making*, 17(03), 883-910.
- Miao, Y. L., Cheng, W. F., Ji, Y. C., Zhang, S., & Kong, Y. L. (2021). Aspect-based sentiment analysis in Chinese based on mobile reviews for BiLSTM-CRF. *Journal of Intelligent & Fuzzy Systems*, 40(5), 8697-8707.
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26.
- Mikolov, T., Yih, W.-t., & Zweig, G. (2013). *Linguistic regularities in continuous space word representations*. Paper presented at the Proceedings of the 2013 conference of the north american chapter of the association for computational linguistics: Human language technologies.
- Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, 11(1), 81. doi:10.1007/s13278-021-00776-6
- Pennington, J., Socher, R., & Manning, C. D. (2014). *Glove: Global vectors for word representation*. Paper presented at the Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP).
- Qiao, B., Zou, Z., Huang, Y., Fang, K., Zhu, X., & Chen, Y. (2022). A joint model for entity and relation extraction based on BERT. *Neural Computing and Applications*, 1-11.

- Rhanoui, M., Mikram, M., Yousfi, S., & Barzali, S. (2019). A CNN-BiLSTM model for document-level sentiment analysis. *Machine Learning and Knowledge Extraction*, 1(3), 832-847.
- Sanderson, M. (2010). Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütze, Introduction to Information Retrieval, Cambridge University Press. 2008. ISBN-13 978-0-521-86571-5, xxi+ 482 pages. *Natural Language Engineering*, 16(1), 100-103.
- Scherer, K. R. (1984). Emotion as a multicomponent process: A model and some cross-cultural data. *Review of personality & social psychology*.
- Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A. Y., & Potts, C. (2013). *Recursive deep models for semantic compositionality over a sentiment treebank*. Paper presented at the Proceedings of the 2013 conference on empirical methods in natural language processing.
- Staliūnaitė, I., & Iacobacci, I. (2020). Compositional and lexical semantics in RoBERTa, BERT and DistilBERT: A case study on CoQA. *arXiv preprint arXiv:2009.08257*.
- Xie, J., Yang, Z., Neubig, G., Smith, N. A., & Carbonell, J. (2018). Neural cross-lingual named entity recognition with minimal resources. *arXiv preprint arXiv:1808.09861*.
- Xin, J., Tang, R., Yu, Y., & Lin, J. (2021). *BERxiT: Early exiting for BERT with better fine-tuning and extension to regression*. Paper presented at the Proceedings of the 16th conference of the European chapter of the association for computational linguistics: Main Volume.
- Xing, F. Z., Cambria, E., & Welsch, R. E. (2018). Natural language based financial forecasting: a survey. *Artificial Intelligence Review*, 50(1), 49-73. doi:10.1007/s10462-017-9588-9
- Xu, G., Meng, Y., Qiu, X., Yu, Z., & Wu, X. (2019). Sentiment analysis of comment texts based on BiLSTM. *Ieee Access*, 7, 51522-51532.
- Yang, L., Li, Y., Wang, J., & Sherratt, R. S. (2020). Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. *IEEE access*, 8, 23522-23530.
- Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R., & Le, Q. V. (2019). Xlnet: Generalized autoregressive pretraining for language understanding. *Advances in neural information processing systems*, 32.
- Yoon, J., & Kim, H. (2017). *Multi-channel lexicon integrated CNN-BiLSTM models for sentiment analysis*. Paper presented at the Proceedings of the 29th conference on computational linguistics and speech processing (ROCLING 2017).
- Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4), e1253.
- Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q., & Artzi, Y. (2019). Bertscore: Evaluating text generation with bert. *arXiv preprint arXiv:1904.09675*.
- Zhang, Y., He, R., Liu, Z., Lim, K. H., & Bing, L. (2020). An unsupervised sentence embedding method by mutual information maximization. *arXiv preprint arXiv:2009.12061*.