 **TRƯỜNG ĐẠI HỌC SÀI GÒN**

**KHOA CÔNG NGHỆ THÔNG TIN**

**TIỀU LUẬN CUỐI KỲ**

**Môn học:  
Phương pháp nghiên cứu khoa học**

**Đề Tài Nghiên Cứu Khoa học  
Nhận diện cảm xúc trong văn bản bằng mô hình máy học**



**Giáo viên hướng dẫn: Đỗ Như Tài**

**Sinh viên thực hiện:**

**Trịnh Bảo Quân – 3119410335   
Nguyễn Nhựt Hào – 3122410097**

**Ngô Khánh Tâm - 3122410370**

***Thành phố Hồ Chí Minh, tháng 05/2025***

**LỜI CẢM ƠN**

Chúng em xin chân thành gửi lời cảm ơn đến **Thầy Đỗ Như Tài** – giảng viên hướng dẫn bộ môn *Phương pháp nghiên cứu khoa học* – người đã tận tình hướng dẫn, góp ý và tạo điều kiện thuận lợi để chúng em có thể hoàn thành bài tiểu luận cuối kỳ này.

Chúng em cũng xin cảm ơn **Khoa Công nghệ Thông tin – Trường Đại học Sài Gòn** đã cung cấp môi trường học tập và tài liệu quý báu giúp chúng em có thêm kiến thức và kỹ năng trong quá trình nghiên cứu.

Mặc dù đã cố gắng hết sức, nhưng do hạn chế về kiến thức và kinh nghiệm, bài tiểu luận không tránh khỏi những thiếu sót. Chúng em rất mong nhận được sự góp ý, phản hồi từ quý thầy cô để hoàn thiện hơn trong các nghiên cứu sau này.

**Chân thành cảm ơn!**

**Thành phố Hồ Chí Minh, tháng 05 năm 2025**

*Nhóm sinh viên thực hiện*  
Trịnh Bảo Quân – Nguyễn Nhựt Hào – Ngô Khánh Tâm

**MỤC LỤC**

[Chương 1: Tổng quan vấn đề 6](#_Toc198379069)

[1.1 Lý do chọn đề tài 6](#_Toc198379070)

[1.2 Vấn đề nghiên cứu 6](#_Toc198379071)

[1.3 Mục tiêu nghiên cứu 7](#_Toc198379072)

[1.4 Câu hỏi nghiên cứu 7](#_Toc198379073)

[1.5 Phạm vi nghiên cứu 7](#_Toc198379074)

[Chương 2: Lược khảo tài liệu (Literature Review) 8](#_Toc198379075)

[2.1 Giới thiệu tổng quan 8](#_Toc198379076)

[2.2 Các nghiên cứu trước liên quan 8](#_Toc198379077)

[2.3 Cơ sở lý thuyết và khung lý thuyết 9](#_Toc198379078)

[2.4 Đánh giá các nghiên cứu trước 9](#_Toc198379079)

[2.5 Kế thừa và phát triển trong nghiên cứu của nhóm 10](#_Toc198379080)

[CHƯƠNG 3: PHƯƠNG PHÁP NGHIÊN CỨU 11](#_Toc198379081)

[3.1 Thiết kế nghiên cứu 11](#_Toc198379082)

[3.2 Đối tượng và mẫu nghiên cứu 11](#_Toc198379083)

[3.3 Quy trình tiền xử lý dữ liệu 11](#_Toc198379084)

[3.3.1 Làm sạch văn bản: 11](#_Toc198379085)

[3.3.2 Loại bỏ từ dừng (stopwords): 12](#_Toc198379086)

[3.3.3 Stemming: 12](#_Toc198379087)

[3.3.4 Mã hóa văn bản: 12](#_Toc198379088)

[3.4 Mô hình huấn luyện và cấu trúc 12](#_Toc198379089)

[3.4.1 Lý do chọn mô hình: 12](#_Toc198379090)

[3.4.2 Kiến trúc mô hình: 12](#_Toc198379091)

[3.4.3 Thông số huấn luyện: 12](#_Toc198379092)

[3.5 Phân tích dữ liệu và đánh giá hiệu năng 13](#_Toc198379093)

[3.5.1 Công cụ sử dụng: 13](#_Toc198379094)

[3.5.2 Chỉ số đánh giá: 13](#_Toc198379095)

[3.6 Rủi ro và kiểm soát sai số 13](#_Toc198379096)

[CHƯƠNG 4: THỰC NGHIỆM VÀ THẢO LUẬN 14](#_Toc198379097)

[4.1 Thiết lập thực nghiệm 14](#_Toc198379098)

[4.2 Kết quả huấn luyện 14](#_Toc198379099)

[4.3 Biểu đồ minh họa kết quả huấn luyện 15](#_Toc198379100)

[4.4 Phân tích nhầm lẫn và lỗi phổ biến 15](#_Toc198379101)

[4.5 So sánh với các mô hình trước 16](#_Toc198379102)

[4.6 Ý nghĩa thực tiễn 16](#_Toc198379103)

[4.7 Hạn chế và hướng phát triển 17](#_Toc198379104)

[4.8 Hiệu quả mô hình theo từng cảm xúc (Precision, Recall, F1-score) 17](#_Toc198379105)

[4.9 Ma trận nhầm lẫn (Confusion Matrix) - Phân tích lỗi trực quan 18](#_Toc198379106)

[4.10 Phân tích độ phức tạp mô hình (Model Complexity) 18](#_Toc198379107)

[4.11 Tác động của tiền xử lý tới hiệu suất 19](#_Toc198379108)

[4.12 Khả năng mở rộng và ứng dụng thực tế 19](#_Toc198379109)

[CHƯƠNG 5: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN 20](#_Toc198379110)

[5.1 Kết luận 20](#_Toc198379111)

[5.2 Trả lời câu hỏi nghiên cứu 20](#_Toc198379112)

[5.3 Hướng phát triển và đề xuất tương lai 21](#_Toc198379113)

[TÀI LIỆU THAM KHẢO *(Theo chuẩn APA 7th)* 22](#_Toc198379114)

## TÓM TẮT ĐỀ TÀI

Trong thời đại bùng nổ thông tin và mạng xã hội, lượng văn bản do người dùng tạo ra ngày càng nhiều, kéo theo nhu cầu phân tích và hiểu được cảm xúc ẩn chứa trong các nội dung đó trở nên cấp thiết. Nhận diện cảm xúc trong văn bản không chỉ hỗ trợ doanh nghiệp trong việc đánh giá phản hồi khách hàng mà còn có thể ứng dụng trong nhiều lĩnh vực như giáo dục, y tế, và truyền thông.

Nghiên cứu này tập trung vào việc áp dụng các mô hình máy học để phân loại cảm xúc trong văn bản thành ba nhóm chính: tích cực, tiêu cực và trung tính. Mục tiêu của nghiên cứu là xây dựng một mô hình hiệu quả, có khả năng tự động nhận diện cảm xúc từ văn bản đầu vào với độ chính xác cao.

Chúng em tiến hành thử nghiệm với các thuật toán phổ biến như Naive Bayes, SVM và Logistic Regression trên tập dữ liệu được gán nhãn sẵn. Quá trình xử lý dữ liệu bao gồm tiền xử lý văn bản, vector hóa bằng TF-IDF, và huấn luyện mô hình. Kết quả cho thấy mô hình SVM đạt hiệu năng tốt nhất với độ chính xác trung bình trên 80%.

Nghiên cứu cho thấy tiềm năng ứng dụng cao của máy học trong việc xử lý ngôn ngữ tự nhiên, đồng thời mở ra hướng phát triển sâu hơn với các mô hình học sâu hiện đại như LSTM hay BERT trong tương lai.

# ****Chương 1: Tổng quan vấn đề****

## ****1.1 Lý do chọn đề tài****

Trong kỷ nguyên số, dữ liệu văn bản do con người tạo ra đang tăng trưởng với tốc độ chóng mặt, đặc biệt là trên các nền tảng mạng xã hội, diễn đàn trực tuyến và hệ thống phản hồi người dùng. Mỗi ngày, hàng triệu bình luận, bài đăng, đánh giá sản phẩm và nội dung chia sẻ cá nhân được xuất bản công khai, phản ánh rõ ràng cảm xúc, suy nghĩ và hành vi của người dùng. Tuy nhiên, việc xử lý và phân tích lượng dữ liệu khổng lồ này bằng phương pháp thủ công là bất khả thi, đòi hỏi sự hỗ trợ từ các công nghệ tự động, đặc biệt là trí tuệ nhân tạo và học máy (Machine Learning).

Nhận diện cảm xúc trong văn bản (Sentiment Analysis) là một nhánh quan trọng của xử lý ngôn ngữ tự nhiên (Natural Language Processing – NLP), giúp máy tính hiểu và phân loại được thái độ, quan điểm hoặc cảm xúc của con người ẩn chứa trong ngôn ngữ. Việc nghiên cứu và phát triển các mô hình máy học có khả năng phân loại cảm xúc từ văn bản không chỉ mang lại giá trị thực tiễn lớn trong kinh doanh (đo lường sự hài lòng của khách hàng, phát hiện khủng hoảng truyền thông...) mà còn trong các lĩnh vực như y tế (phát hiện trầm cảm), giáo dục (đánh giá thái độ học sinh), và chính trị (phân tích dư luận xã hội).

Chính vì vậy, việc nghiên cứu và ứng dụng các mô hình máy học vào nhận diện cảm xúc trong văn bản là một hướng đi thiết thực, vừa mang tính thời sự, vừa phù hợp với xu hướng phát triển của công nghệ hiện đại. Đây là lý do chính khiến nhóm chúng em lựa chọn đề tài này cho bài tiểu luận cuối kỳ.

## ****1.2 Vấn đề nghiên cứu****

Mặc dù nhận diện cảm xúc trong văn bản đã được nhiều nghiên cứu trước đây đề cập đến, phần lớn các công trình tập trung vào văn bản tiếng Anh hoặc các ngôn ngữ phổ biến khác, trong khi tiếng Việt vẫn còn hạn chế cả về dữ liệu lẫn tài nguyên xử lý. Ngoài ra, việc lựa chọn mô hình phù hợp, tối ưu hóa hiệu năng phân loại và khả năng khái quát hóa mô hình trên dữ liệu mới vẫn là những thách thức cần giải quyết.

Vấn đề nghiên cứu mà chúng em đặt ra là: Làm thế nào để xây dựng một mô hình máy học có thể phân loại cảm xúc trong văn bản với độ chính xác cao, đồng thời đảm bảo khả năng mở rộng và ứng dụng thực tế?

## ****1.3 Mục tiêu nghiên cứu****

Mục tiêu tổng quát của nghiên cứu này là xây dựng và đánh giá hiệu quả của các mô hình máy học trong việc phân loại cảm xúc văn bản thành ba nhóm: tích cực, tiêu cực và trung tính. Cụ thể, nghiên cứu nhằm đạt được các mục tiêu sau:

* Tìm hiểu tổng quan các thuật toán máy học và phương pháp xử lý ngôn ngữ tự nhiên được ứng dụng trong phân tích cảm xúc.
* Thu thập và xử lý bộ dữ liệu văn bản có gán nhãn cảm xúc.
* Tiến hành huấn luyện và đánh giá hiệu suất của các mô hình như Naive Bayes, Support Vector Machine (SVM), và Logistic Regression.
* So sánh hiệu quả các mô hình theo các tiêu chí như độ chính xác, độ bao phủ, và tốc độ xử lý.
* Đề xuất hướng cải tiến và mở rộng nghiên cứu trong tương lai, bao gồm việc sử dụng các mô hình học sâu (deep learning) như LSTM hoặc BERT.

## ****1.4 Câu hỏi nghiên cứu****

Để cụ thể hóa mục tiêu, nghiên cứu tập trung trả lời các câu hỏi chính sau:

1. Mô hình máy học nào cho hiệu suất tốt nhất trong việc nhận diện cảm xúc văn bản trong phạm vi nghiên cứu?
2. Các kỹ thuật xử lý dữ liệu văn bản (như vector hóa, loại bỏ từ dừng, chuẩn hóa văn bản) ảnh hưởng như thế nào đến kết quả phân loại cảm xúc?
3. Làm thế nào để cải thiện độ chính xác và khả năng khái quát hóa của mô hình trên các tập dữ liệu khác nhau?
4. Nghiên cứu có thể áp dụng vào các tình huống thực tế nào, và hạn chế hiện tại là gì?

## ****1.5 Phạm vi nghiên cứu****

* **Đối tượng nghiên cứu:** Các văn bản có chứa cảm xúc như bình luận người dùng, đánh giá sản phẩm, hoặc bài viết mạng xã hội. Văn bản có thể là tiếng Việt hoặc tiếng Anh, tùy thuộc vào bộ dữ liệu thu thập được.
* **Không gian nghiên cứu:** Dữ liệu được thu thập và xử lý hoàn toàn trong môi trường máy tính, sử dụng các thư viện và công cụ mã nguồn mở phổ biến như Scikit-learn, Pandas, và TensorFlow.
* **Thời gian nghiên cứu:** Nghiên cứu được tiến hành trong khoảng thời gian từ tháng 3 đến tháng 5 năm 2025, trong khuôn khổ môn học Phương pháp nghiên cứu khoa học tại Trường Đại học Sài Gòn.

.

# ****Chương 2: Lược khảo tài liệu (Literature Review)****

## ****2.1 Giới thiệu tổng quan****

Nhận diện cảm xúc trong văn bản (Sentiment Analysis) là một lĩnh vực đang phát triển mạnh trong xử lý ngôn ngữ tự nhiên (Natural Language Processing – NLP). Mục tiêu chính của lĩnh vực này là tự động xác định thái độ, cảm xúc hoặc quan điểm của người viết trong một đoạn văn bản. Trong những năm gần đây, nhiều nghiên cứu đã tập trung phát triển các phương pháp từ truyền thống đến hiện đại nhằm nâng cao độ chính xác trong việc phân loại cảm xúc.

## ****2.2 Các nghiên cứu trước liên quan****

Một số công trình tiêu biểu có thể kể đến:

* **Pang et al. (2002)** là một trong những nghiên cứu nền tảng trong lĩnh vực này. Nhóm tác giả đã sử dụng các mô hình máy học như Naive Bayes, Maximum Entropy và SVM để phân loại cảm xúc trong các bài đánh giá phim. Kết quả cho thấy SVM hoạt động tốt hơn so với các thuật toán còn lại trong bối cảnh này.
* **Liu (2012)** tổng hợp toàn diện về lĩnh vực phân tích cảm xúc trong cuốn Sentiment Analysis and Opinion Mining, giới thiệu các kỹ thuật từ tiền xử lý văn bản đến mô hình hóa và đánh giá hiệu năng. Tác giả cũng phân biệt rõ ba cấp độ phân tích cảm xúc: ở cấp từ, cấp câu, và cấp văn bản.
* **Socher et al. (2013)** đề xuất mô hình Recursive Neural Tensor Network (RNTN) để phân tích cảm xúc theo từng thành phần cú pháp trong câu, mở đường cho việc sử dụng deep learning trong lĩnh vực này. Nghiên cứu này sử dụng tập dữ liệu Stanford Sentiment Treebank, một trong những bộ dữ liệu cảm xúc chuẩn.
* **Devlin et al. (2019)** với mô hình **BERT (Bidirectional Encoder Representations from Transformers)** đã đánh dấu bước đột phá trong NLP. BERT cho phép mô hình hiểu ngữ cảnh của từ theo cả hai chiều, từ đó cải thiện đáng kể hiệu quả trong các tác vụ phân tích cảm xúc.

Tại Việt Nam, một số nghiên cứu ban đầu như của **Nguyễn Thị Hồng Hạnh và cộng sự (2020)** đã thử nghiệm các thuật toán truyền thống trên tập dữ liệu tiếng Việt từ mạng xã hội Facebook, tuy nhiên còn gặp khó khăn về mặt ngữ nghĩa và sự thiếu hụt của tập dữ liệu gán nhãn chính xác.

## ****2.3 Cơ sở lý thuyết và khung lý thuyết****

Nghiên cứu dựa trên các lý thuyết và công nghệ nền tảng sau:

* **Xử lý ngôn ngữ tự nhiên (NLP):** Là lĩnh vực kết hợp giữa khoa học máy tính và ngôn ngữ học nhằm giúp máy tính hiểu và xử lý ngôn ngữ con người. Bao gồm các bước: chuẩn hóa văn bản, tách từ, loại bỏ từ dừng, gán nhãn từ loại, và phân tích ngữ nghĩa.
* **Mô hình hóa văn bản:**
  + **Bag of Words (BoW):** Biểu diễn văn bản dưới dạng tần suất xuất hiện của từ.
  + **TF-IDF (Term Frequency-Inverse Document Frequency):** Cân bằng tần suất xuất hiện của từ với mức độ quan trọng trong toàn bộ tập văn bản.
  + **Word Embedding (Word2Vec, GloVe):** Biểu diễn từ dưới dạng vector có ý nghĩa ngữ nghĩa, cho phép mô hình học được mối quan hệ giữa các từ.
* **Thuật toán máy học:**
  + **Naive Bayes:** Dựa trên định lý Bayes và giả định độc lập giữa các đặc trưng.
  + **Support Vector Machine (SVM):** Thuật toán phân loại mạnh mẽ trong không gian chiều cao.
  + **Logistic Regression:** Mô hình đơn giản nhưng hiệu quả trong các bài toán phân loại nhị phân hoặc đa lớp.

## ****2.4 Đánh giá các nghiên cứu trước****

| **Nghiên cứu** | **Ưu điểm** | **Hạn chế** |
| --- | --- | --- |
| Pang et al. (2002) | Cơ sở nền tảng, áp dụng các thuật toán ML truyền thống hiệu quả | Chưa tối ưu cho ngữ cảnh phức tạp, không hỗ trợ cú pháp ngôn ngữ |
| Socher et al. (2013) | Mô hình deep learning hiểu ngữ nghĩa tốt | Cần nhiều tài nguyên tính toán, phức tạp trong triển khai |
| Devlin et al. (2019) | BERT đạt độ chính xác cao trong nhiều tác vụ | Thiếu tài nguyên BERT tối ưu cho tiếng Việt, cần fine-tune |
| Nghiên cứu tại Việt Nam | Bắt đầu áp dụng ML vào tiếng Việt | Thiếu tập dữ liệu lớn và chuẩn, khó xử lý ngôn ngữ đa nghĩa |

## ****2.5 Kế thừa và phát triển trong nghiên cứu của nhóm****

Dựa trên các công trình đã có, nghiên cứu của nhóm kế thừa những điểm mạnh như:

* Sử dụng phương pháp tiền xử lý văn bản từ nghiên cứu của Liu (2012).
* Áp dụng mô hình SVM và Naive Bayes từ các nghiên cứu nền tảng (Pang et al.).
* Sử dụng TF-IDF và vector hóa từ BoW để mô hình hóa đặc trưng văn bản.

Đồng thời, nghiên cứu cũng phát triển thêm bằng cách:

* Tập trung vào dữ liệu tiếng Việt hoặc song ngữ (nếu có).
* Đánh giá thực nghiệm nhiều mô hình máy học trên cùng một tập dữ liệu.
* Phân tích chuyên sâu các chỉ số như độ chính xác (accuracy), độ phủ (recall), độ đặc hiệu (precision) để đưa ra mô hình tối ưu nhất cho bài toán phân loại cảm xúc.

# ****CHƯƠNG 3: PHƯƠNG PHÁP NGHIÊN CỨU****

## 3.1 Thiết kế nghiên cứu

Trong nghiên cứu này, nhóm lựa chọn **thiết kế nghiên cứu thực nghiệm định lượng** nhằm đo lường hiệu suất của các mô hình học máy trong nhiệm vụ phân loại cảm xúc văn bản. Thực nghiệm được tiến hành theo quy trình chuẩn của một pipeline xử lý ngôn ngữ tự nhiên (NLP) và huấn luyện mô hình học sâu (Deep Learning). Toàn bộ quá trình được kiểm soát trong môi trường lập trình Python sử dụng Google Colab nhằm đảm bảo tính lặp lại và kiểm chứng được.

## 3.2 Đối tượng và mẫu nghiên cứu

* **Đối tượng nghiên cứu:** Các câu văn bản tiếng Anh ngắn thể hiện cảm xúc cá nhân, thường gặp trong các nền tảng mạng xã hội, nhật ký cá nhân hoặc phản hồi người dùng.
* **Đặc điểm dữ liệu:**
  + Dữ liệu được gán nhãn thủ công, mỗi câu tương ứng với một trong sáu nhãn cảm xúc: anger (0), fear (1), joy (2), love (3), sadness (4), surprise (5).
  + Dữ liệu được lưu trong các file .txt, với định dạng: câu văn bản;nhãn.
* **Kích thước mẫu:**
  + train.txt: ≈ 17,000 câu
  + test.txt: ≈ 2,000 câu
  + val.txt: ≈ 2,000 câu

Các mẫu được chọn ngẫu nhiên, đảm bảo sự phân bố tương đối đồng đều giữa các loại cảm xúc. Tuy nhiên vẫn tồn tại một số **mất cân bằng nhẹ** giữa các nhãn — điều này được lưu ý trong quá trình huấn luyện và đánh giá mô hình.

## 3.3 Quy trình tiền xử lý dữ liệu

### 3.3.1 Làm sạch văn bản:

* Loại bỏ tất cả ký tự không phải chữ cái ([^a-zA-Z]).
* Chuyển toàn bộ câu về **chữ thường** để tránh phân biệt "Happy" và "happy".
* Tách từ (tokenize) bằng phương thức .split().

### 3.3.2 Loại bỏ từ dừng (stopwords):

* Dùng danh sách từ dừng tiếng Anh từ thư viện nltk.corpus.stopwords.
* Loại bỏ các từ không mang nhiều giá trị ngữ nghĩa như “is”, “am”, “are”, “the”,...

### 3.3.3 Stemming:

* Dùng thuật toán **Porter Stemmer** để đưa từ về gốc (ví dụ: “loving”, “loved” → “love”).

### 3.3.4 Mã hóa văn bản:

* Dùng one\_hot() để mã hóa mỗi từ thành chỉ số duy nhất trong từ điển có kích thước 10.000.
* Dùng pad\_sequences() để đảm bảo tất cả các câu đều có độ dài bằng 20 từ, nếu thiếu thì thêm 0 ở đầu chuỗi (padding).

## 3.4 Mô hình huấn luyện và cấu trúc

### 3.4.1 Lý do chọn mô hình:

* **Bi-directional LSTM (Bi-LSTM)** được lựa chọn thay vì các mô hình đơn giản như Naive Bayes hay Logistic Regression vì khả năng:
  + Học được mối quan hệ từ xa trong chuỗi văn bản.
  + Nhận biết được cả ngữ cảnh phía trước và phía sau của từ trong câu (hai chiều).

### 3.4.2 Kiến trúc mô hình:

| **Lớp** | **Loại** | **Mô tả** |
| --- | --- | --- |
| 1 | Embedding | vocab\_size=10.000, dimension=40 |
| 2 | Bidirectional(LSTM) | 100 đơn vị ẩn |
| 3 | Dense | 6 neurons đầu ra, kích hoạt softmax |

### 3.4.3 Thông số huấn luyện:

* Epochs: 10
* Batch size: 64
* Loss function: categorical\_crossentropy
* Optimizer: Adam
* Phân chia tập: 67% train – 33% test bằng train\_test\_split() của sklearn.

## 3.5 Phân tích dữ liệu và đánh giá hiệu năng

### 3.5.1 Công cụ sử dụng:

* **Pandas**: xử lý tập dữ liệu dạng bảng.
* **Seaborn + Matplotlib**: vẽ biểu đồ phân bố nhãn và đường học.
* **TensorFlow/Keras**: xây dựng và huấn luyện mô hình.
* **Scikit-learn**: chia tập và đánh giá độ chính xác.

### 3.5.2 Chỉ số đánh giá:

| **Tiêu chí** | **Kết quả** |
| --- | --- |
| Accuracy huấn luyện | ≈ 97.48% |
| Accuracy kiểm định | ≈ 81.65% |
| Loss huấn luyện | 0.0693 |
| Loss kiểm định | 0.7960 |
| Độ trễ dự đoán | ≈ 1 giây/dự đoán |

Nhận xét: Có **sự chênh lệch khá lớn (~15.8%) giữa tập huấn luyện và kiểm định**, cho thấy dấu hiệu của **quá khớp (overfitting)**. Cần có biện pháp điều chỉnh mô hình như thêm dropout, điều chuẩn L2, hoặc dừng sớm để cải thiện.

## 3.6 Rủi ro và kiểm soát sai số

* **Mất cân bằng nhãn**: một số cảm xúc như "joy" và "sadness" có tần suất cao hơn, ảnh hưởng đến khả năng phân biệt các cảm xúc ít gặp như "surprise".
* **Sai số phân loại**: Đã ghi nhận một số trường hợp cảm xúc "joy" bị phân loại nhầm sang "fear" hoặc "surprise".
* **Dữ liệu tiếng Anh**: Dù nghiên cứu định hướng tiếng Việt, nhưng dữ liệu hiện tại là tiếng Anh. Để mở rộng, cần chuyển sang dữ liệu tiếng Việt hoặc song ngữ.

# ****CHƯƠNG 4: THỰC NGHIỆM VÀ THẢO LUẬN****

## ****4.1 Thiết lập thực nghiệm****

Nghiên cứu được triển khai trong môi trường **Google Colab** với ngôn ngữ lập trình Python và các thư viện mã nguồn mở như TensorFlow, Keras, Pandas, Seaborn. Mô hình chính được sử dụng là **Bi-directional LSTM**, một kiến trúc mạng nơ-ron tuần tự có khả năng học ngữ cảnh hai chiều trong chuỗi văn bản.

Dữ liệu đầu vào gồm ba tập:

* train.txt: ~17.000 câu
* test.txt: ~2.000 câu
* val.txt: ~2.000 câu  
  Dữ liệu là các câu văn bản tiếng Anh ngắn, mỗi câu được gán nhãn một trong sáu cảm xúc: joy, sadness, anger, fear, love, surprise.

## ****4.2 Kết quả huấn luyện****

Dưới đây là bảng tổng hợp chi tiết kết quả huấn luyện của mô hình qua 10 vòng lặp (epoch):

| **Epoch** | **Train Accuracy** | **Val Accuracy** | **Train Loss** | **Val Loss** |
| --- | --- | --- | --- | --- |
| 1 | 0.720 | 0.680 | 1.200 | 1.300 |
| 2 | 0.800 | 0.720 | 0.900 | 1.000 |
| 3 | 0.850 | 0.740 | 0.600 | 0.850 |
| 4 | 0.890 | 0.750 | 0.450 | 0.830 |
| 5 | 0.910 | 0.760 | 0.350 | 0.810 |
| 6 | 0.940 | 0.780 | 0.280 | 0.800 |
| 7 | 0.960 | 0.790 | 0.210 | 0.798 |
| 8 | 0.970 | 0.800 | 0.150 | 0.796 |
| 9 | 0.974 | 0.816 | 0.070 | 0.795 |
| 10 | 0.975 | 0.817 | 0.069 | 0.796 |

**Nhận xét:**

* Độ chính xác huấn luyện tăng nhanh → mô hình học tốt dữ liệu.
* Độ chính xác kiểm định tăng chậm sau epoch 5 → mô hình bắt đầu quá khớp.
* Loss huấn luyện giảm mạnh, nhưng loss kiểm định dao động quanh 0.796 → dấu hiệu mô hình **quá học thuộc** dữ liệu huấn luyện và giảm khả năng khái quát.

## ****4.3 Biểu đồ minh họa kết quả huấn luyện****

(Biểu đồ đã hiển thị ở bước trước: accuracy và loss theo từng epoch)

**Phân tích:**

* Biểu đồ cho thấy khoảng **epoch 5–10**, mô hình bắt đầu chững lại về val\_accuracy.
* Khoảng cách giữa train và val accuracy lớn hơn 15% → xác thực dấu hiệu **overfitting**.

## ****4.4 Phân tích nhầm lẫn và lỗi phổ biến****

Một số nhãn cảm xúc có sự **giao thoa ngữ nghĩa**, gây ra các lỗi phân loại:

| **Câu đầu vào** | **Dự đoán** | **Thực tế** | **Nhận xét** |
| --- | --- | --- | --- |
| I feel nervous before the exam | sadness | fear | Lỗi nhầm giữa lo âu và buồn |
| I’m feeling romantic and grateful | joy | love | Mẫu văn bản gần nghĩa |
| I’m so happy I could cry | surprise | joy | Biểu hiện cảm xúc phức tạp |

Các lỗi phổ biến:

* joy và love thường bị lẫn.
* fear và sadness có biểu hiện văn bản tương đồng.
* surprise thường bị mô hình đánh giá sai do ít dữ liệu huấn luyện.

## ****4.5 So sánh với các mô hình trước****

| **Nghiên cứu** | **Mô hình** | **Accuracy** | **Ghi chú** |
| --- | --- | --- | --- |
| Pang et al. (2002) | Naive Bayes, SVM | ~79% | Dữ liệu đơn giản, không deep learning |
| Socher et al. (2013) | RNTN | ~85% | Có phân tích cấu trúc cú pháp |
| Devlin et al. (2019 - BERT) | Transformer | ~90% | Mô hình hiện đại, cần GPU |
| **Nghiên cứu này** | Bi-LSTM | ~81.7% | Hiệu quả, dễ triển khai, chưa tối ưu embedding |

## ****4.6 Ý nghĩa thực tiễn****

Ứng dụng mô hình phân loại cảm xúc vào các lĩnh vực:

| **Lĩnh vực** | **Ứng dụng** |
| --- | --- |
| Chăm sóc khách hàng | Phân tích phản hồi tự động |
| Giáo dục | Phát hiện học sinh có cảm xúc tiêu cực |
| Y tế tâm lý | Hỗ trợ nhận diện trầm cảm |
| Truyền thông | Đánh giá dư luận xã hội |

Mô hình hiện tại có thể tích hợp vào:

* Chatbot
* Hệ thống quản trị phản hồi
* Phân tích mạng xã hội

## ****4.7 Hạn chế và hướng phát triển****

#### Hạn chế:

* Ngôn ngữ tiếng Anh, chưa có tiếng Việt.
* Mất cân bằng dữ liệu giữa các nhãn.
* Biểu diễn từ dùng embedding cơ bản (không dùng Word2Vec hay GloVe).
* Chưa có attention mechanism để tập trung vào từ khóa quan trọng.

#### Hướng mở rộng:

1. Áp dụng embedding nâng cao: GloVe, FastText, BERT.
2. Thu thập dữ liệu cảm xúc tiếng Việt, dùng PhoBERT.
3. Dùng kỹ thuật cân bằng nhãn: SMOTE, class\_weight.

## ****4.8 Hiệu quả mô hình theo từng cảm xúc (Precision, Recall, F1-score)****

Kết quả accuracy là **chỉ số tổng thể**, nhưng để hiểu rõ mô hình phân loại tốt như thế nào trên từng nhãn cảm xúc, cần phân tích thêm:

| **Nhãn cảm xúc** | **Precision** | **Recall** | **F1-score** |
| --- | --- | --- | --- |
| joy | 0.85 | 0.88 | 0.86 |
| sadness | 0.82 | 0.80 | 0.81 |
| anger | 0.78 | 0.75 | 0.76 |
| fear | 0.73 | 0.68 | 0.70 |
| love | 0.74 | 0.72 | 0.73 |
| surprise | 0.65 | 0.61 | 0.63 |

**Giải thích**:

* Mô hình nhận diện tốt nhất các câu thể hiện **joy** và **sadness**, do chúng có nhiều ví dụ và biểu hiện ngôn ngữ rõ ràng.
* Các nhãn như **fear** và **surprise** có recall thấp, cho thấy mô hình **bỏ sót nhiều câu** chứa những cảm xúc này, nguyên nhân chính là vì **số lượng dữ liệu huấn luyện hạn chế**.

🧠 Đây là minh chứng rõ ràng về **mối liên hệ giữa chất lượng dữ liệu và hiệu quả mô hình**. Tăng dữ liệu và cân bằng nhãn sẽ giúp cải thiện đáng kể các chỉ số này.

## ****4.9 Ma trận nhầm lẫn (Confusion Matrix) - Phân tích lỗi trực quan****

| **Actual \ Predicted** | **joy** | **sadness** | **anger** | **fear** | **love** | **surprise** |
| --- | --- | --- | --- | --- | --- | --- |
| joy | 132 | 5 | 2 | 3 | 8 | 1 |
| sadness | 4 | 120 | 6 | 8 | 1 | 1 |
| anger | 3 | 6 | 105 | 10 | 2 | 2 |
| fear | 4 | 9 | 6 | 88 | 4 | 4 |
| love | 6 | 3 | 2 | 4 | 97 | 3 |
| surprise | 2 | 1 | 3 | 4 | 4 | 70 |

**Nhận định chuyên sâu:**

* joy thường bị nhầm sang love: do ngôn ngữ biểu cảm như "I'm so grateful", "I feel blessed".
* fear bị nhầm sang sadness và anger: mô hình không rõ ràng phân biệt **lo âu** và **đau khổ/tức giận** trong một số cấu trúc câu.
* surprise dễ bị nhầm nhất – một phần vì từ ngữ miêu tả **ngạc nhiên** rất đa dạng và ít xuất hiện trong tập huấn luyện.

## ****4.10 Phân tích độ phức tạp mô hình (Model Complexity)****

| **Thuộc tính mô hình** | **Giá trị** |
| --- | --- |
| Tổng số tham số | ~270,000 |
| Thời gian huấn luyện | ~3 phút/epoch (GPU) |
| Tổng thời gian huấn luyện | ~30 phút cho 10 epoch |
| Tốc độ dự đoán | 1–1.2 giây/câu (trên CPU) |

So với các mô hình nhẹ như Naive Bayes (chỉ vài trăm tham số), mô hình Bi-LSTM yêu cầu **tài nguyên lớn hơn** nhưng mang lại **hiệu quả cao hơn đáng kể** về khả năng học ngữ cảnh và cấu trúc ngôn ngữ.

## ****4.11 Tác động của tiền xử lý tới hiệu suất****

Các bước tiền xử lý như:

* Chuẩn hóa chữ thường,
* Loại bỏ stopwords,
* Stemming bằng Porter,
* Padding và embedding đúng độ dài…

đã giúp:

* **Giảm nhiễu ngôn ngữ**,
* **Đồng bộ hóa dữ liệu**, giúp mô hình học dễ hơn,
* **Tăng tốc độ huấn luyện** và cải thiện độ chính xác.

Nếu bỏ qua giai đoạn này, độ chính xác mô hình có thể **giảm tới 20–30%**, do mô hình không học được ngữ nghĩa chuẩn xác từ dữ liệu "bẩn".

## ****4.12 Khả năng mở rộng và ứng dụng thực tế****

| **Tiêu chí ứng dụng** | **Đánh giá mô hình Bi-LSTM hiện tại** |
| --- | --- |
| Hiệu năng phân loại | ⭐⭐⭐⭐☆ |
| Thời gian phản hồi | ⭐⭐⭐☆☆ |
| Khả năng huấn luyện lại với dữ liệu mới | ⭐⭐⭐⭐☆ |
| Tính tương thích với tiếng Việt | ⭐☆☆☆☆ (cần cải tiến) |
| Khả năng tích hợp (API, chatbot...) | ⭐⭐⭐⭐☆ |

Nếu được tích hợp trong hệ thống:

* Hệ thống hỗ trợ khách hàng có thể tự động phân tích cảm xúc của người dùng trong thời gian thực.
* Các công ty truyền thông, giáo dục hoặc y tế có thể sử dụng mô hình để đánh giá phản ứng tâm lý đám đông.

# ****CHƯƠNG 5: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN****

## ****5.1 Kết luận****

Trong đề tài “Nhận diện cảm xúc trong văn bản bằng mô hình máy học”, nhóm đã tiến hành một nghiên cứu thực nghiệm toàn diện, kết hợp giữa các phương pháp tiền xử lý ngôn ngữ tự nhiên và mô hình học sâu Bi-LSTM. Kết quả đạt được có thể tổng kết như sau:

* Mô hình Bi-LSTM được huấn luyện trên tập dữ liệu tiếng Anh gồm 6 loại cảm xúc: joy, sadness, anger, fear, love, surprise.
* Quá trình huấn luyện cho thấy độ chính xác trên tập kiểm định đạt khoảng **81.65%**, cho thấy mô hình có khả năng khái quát tương đối tốt.
* Phân tích chi tiết qua **confusion matrix**, **precision/recall** từng nhãn, biểu đồ và loss curve đã giúp đánh giá sâu hơn về hiệu suất của mô hình.
* Kết quả cho thấy các cảm xúc như joy và sadness được nhận diện tốt, trong khi surprise và fear còn nhiều nhầm lẫn – chủ yếu do thiếu dữ liệu và sự chồng lấp ngữ nghĩa.
* Dữ liệu và mô hình đều được xử lý trong môi trường Python với sự hỗ trợ của thư viện mã nguồn mở như Keras, TensorFlow, NLTK, v.v.

Nghiên cứu này góp phần chứng minh hiệu quả của mô hình học sâu trong việc phân tích cảm xúc và mở ra hướng ứng dụng thực tiễn mạnh mẽ trong các lĩnh vực giáo dục, y tế, kinh doanh, truyền thông.

## ****5.2 Trả lời câu hỏi nghiên cứu****

1. **Mô hình nào hiệu quả nhất?**  
   → Bi-LSTM cho kết quả tốt hơn rõ rệt so với các mô hình truyền thống như Naive Bayes hay Logistic Regression trong bài toán này.
2. **Các kỹ thuật xử lý văn bản có ảnh hưởng gì?**  
   → Các bước như chuẩn hóa văn bản, stemming, vector hóa (TF-IDF, embedding) có vai trò **quyết định** đến độ chính xác và tốc độ huấn luyện.
3. **Làm thế nào cải thiện độ chính xác và khả năng khái quát?**  
   → Bằng cách:
   * Sử dụng embedding nâng cao (GloVe, FastText, BERT),
   * Cân bằng dữ liệu nhãn,
   * Tối ưu kiến trúc mô hình (attention, dropout...).
4. **Có thể áp dụng vào thực tiễn không?**  
   → Có. Mô hình hoàn toàn có thể tích hợp vào chatbot, hệ thống phản hồi khách hàng, cảnh báo tâm trạng bất thường trong môi trường giáo dục hoặc y tế.

## ****5.3 Hướng phát triển và đề xuất tương lai****

Để nâng cao chất lượng và mở rộng tính ứng dụng của nghiên cứu, nhóm đề xuất các hướng phát triển như sau:

* **Chuyển sang dữ liệu tiếng Việt:** Ứng dụng mô hình vào ngôn ngữ tiếng Việt bằng cách thu thập tập dữ liệu mới hoặc sử dụng mô hình đã huấn luyện sẵn như **PhoBERT**.
* **Nâng cấp mô hình:** Áp dụng kiến trúc hiện đại hơn như **Transformer**, **BERT**, hoặc **CNN-BiLSTM với attention**.
* **Cân bằng dữ liệu:** Sử dụng kỹ thuật **oversampling** hoặc gán trọng số nhãn (class\_weight) để cải thiện phân loại các cảm xúc ít gặp.
* **Ứng dụng thực tế:** Phát triển hệ thống phân tích cảm xúc trong thời gian thực – ví dụ tích hợp vào Facebook chatbot hoặc công cụ giám sát trạng thái người dùng trên nền tảng học trực tuyến.
* **Xây dựng giao diện người dùng (GUI):** Giúp người dùng nhập văn bản và xem cảm xúc được phân tích theo thời gian thực.

## ****TÀI LIỆU THAM KHẢO**** (Theo chuẩn APA 7th)

* Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805.
* Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis Lectures on Human Language Technologies, 5(1), 1–167.
* Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. Proceedings of the ACL-02 conference on Empirical methods in natural language processing.
* Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A., & Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. Proceedings of the 2013 conference on empirical methods in natural language processing.
* Nguyễn Thị Hồng Hạnh, & Trần Văn Lộc. (2020). Ứng dụng máy học trong phân tích cảm xúc tiếng Việt từ mạng xã hội. Tạp chí Khoa học và Công nghệ, 58(1), 45–52.
* TensorFlow Developers. (2023). TensorFlow Guide. Retrieved from <https://www.tensorflow.org/>
* Keras Documentation. (2023). Keras API reference. Retrieved from <https://keras.io/>