|  |  |  |
| --- | --- | --- |
| Logo  Description automatically generated | **TRƯỜNG ĐẠI HỌC KINH TẾ QUỐC DÂN**  **TRƯỜNG CÔNG NGHỆ**  **KHOA CÔNG NGHỆ THÔNG TIN**  \*\*\*\*\*\*\*\*\*\*\*\* |  |

ỨNG DỤNG TRÍ TUỆ NHÂN TẠO TRONG KINH DOANH VÀ QUẢN LÝ

**ThS. Phạm Thảo**

[**Thaop@neu.edu.vn**](mailto:Thaop@neu.edu.vn)

**(Bản thảo**

**Đang cập nhật**

**Lưu hành nội bộ)**

*Hà Nội – 01/2025*

MỤC LỤC

[Chương 1: Hướng dẫn Xây dựng mô hình và phát triển các ứng dụng AI trong kinh doanh và quản lý 4](#_Toc193459087)

[1.1 Mô tả chung 4](#_Toc193459088)

[1.2 Xác định bài toán 4](#_Toc193459089)

[1.2.1 Tên đề tài 4](#_Toc193459090)

[1.2.2 Mục tiêu cùa đề tài 4](#_Toc193459091)

[1.3 Xác Định, Chuẩn hóa dữ liệu 4](#_Toc193459092)

[1.4 Thiết kế và xây dựng mô hình 5](#_Toc193459093)

[1.4.1 Thiết kế 5](#_Toc193459094)

[1.4.2 Xây dựng mô hình 5](#_Toc193459095)

[1.5 Thử nghiệm và đánh giá mô hình 5](#_Toc193459096)

[1.5.1 Làm thế nào để biết mô hình tốt hay chưa tốt? 5](#_Toc193459097)

[1.5.2 Thử và so sánh 6](#_Toc193459098)

[1.5.3 Kết luận 6](#_Toc193459099)

[1.5.4 Xuất mô hình tốt nhất 6](#_Toc193459100)

[1.5.5 Viết một đoạn code để có thể thực thi các trường hợp thử này 6](#_Toc193459101)

[1.6 Xây dựng giao diện người sử dụng 6](#_Toc193459102)

[1.6.1 Xác định đầu vào và đầu ra cho người sử dụng 6](#_Toc193459103)

[1.6.2 Xác định đầu ra 7](#_Toc193459104)

[1.7 Vận hành thử nghiệm 7](#_Toc193459105)

[Chương 2: MỘT TRƯỜNG hợp Phân tích cảm xúc bình luận sản phẩm 8](#_Toc193459106)

[2.1 Mô tả chung 8](#_Toc193459107)

[2.2 Xác định bài toán 8](#_Toc193459108)

[2.2.1 Tên đề tài 8](#_Toc193459109)

[2.2.2 Mục tiêu cùa đề tài 8](#_Toc193459110)

[2.3 Xác Định, Chuẩn hóa dữ liệu 8](#_Toc193459111)

[2.4 Thiết kế và xây dựng mô hình 9](#_Toc193459112)

[2.4.1 Thiết kế 9](#_Toc193459113)

[2.4.2 Xây dựng mô hình 9](#_Toc193459114)

[2.5 Thử nghiệm và đánh giá mô hình 10](#_Toc193459115)

[2.5.1 Làm thế nào để biết mô hình tốt hay chưa tốt? 10](#_Toc193459116)

[2.5.2 Thử và so sánh 10](#_Toc193459117)

[2.5.3 Kết quả thử nghiệm với tập dữ liệu kiểm tra và dữ liệu mới 12](#_Toc193459118)

[2.5.4 Ví dụ Kết quả đào tạo huấn luyện 15](#_Toc193459119)

[2.5.5 Kết luận 15](#_Toc193459120)

[2.5.6 Xuất mô hình tốt nhất 16](#_Toc193459121)

[2.5.7 Viết một đoạn code để có thể thực thi các trường hợp thử này 16](#_Toc193459122)

[2.6 Xây dựng giao diện người sử dụng 16](#_Toc193459123)

[2.6.1 Đề xuất kiến trúc hệ thống 16](#_Toc193459124)

[2.6.2 Xác định đầu vào và đầu ra cho người sử dụng 17](#_Toc193459125)

[2.6.3 Xác định đầu ra 18](#_Toc193459126)

[2.6.4 Ví dụ giao diện người dung 18](#_Toc193459127)

[2.6.5 Quản trị có thể thấy như thế nào 20](#_Toc193459128)

[2.7 Vận hành thử nghiệm 20](#_Toc193459129)

# Hướng dẫn Xây dựng mô hình và phát triển các ứng dụng AI trong kinh doanh và quản lý

## Mô tả chung

Xây dựng một kịch bản, bài toán ứng dụng trong thực tế.

Cần nhấn mạnh giải quyết được một bài toán cụ thể, một tình huống cụ thể trong thực tế mà người làm công việc đó sẽ thực hiện.

Khi chưa có AI thì hoạt động này đang triển khai như thế nào, cụ thể nêu quy trình (bao gồm các bước hoặc ngữ cảnh mà bước đó sẽ thực hiện), từ đó đi sâu vào mô tả chi tiết hoạt động mà mình muốn nghiên cứu và triển khai.

Hình thức của đề tài có thể là một sản phẩm hoàn chỉnh đã được thực thi và triển khai thử nghiệm, nhưng cũng có thể chấp nhận trường hợp chưa lập trình xong nhưng cần được bù đắp và giải thích bằng thực tế tại sao lại chưa triển khai được. Trường hợp này thường là có ý tưởng tốt, nhưng việc lấy dữ liệu có thể triển khai gặp khó khăn, hoặc mô hình cần những điểm mới mà mình đã chỉ ra trong phần thiết kế một cách cụ thể tuy nhiên chưa phát triển xong. Điều này, cần phải được nhấn mạnh trong phần tóm tắt của đề tài nhóm.

## Xác định bài toán

### Tên đề tài

Ví dụ Xét bài toán phân tích cảm xúc từ các bình luận trên hệ thống web thường mại điện tử

Ví dụ Bài viết trên Fanpage về dịch vụ XYZ.

### Mục tiêu cùa đề tài

*Đề tài này nhắm đến mục đích gì:*

Xây dựng được một mô hình ….để thực hiện ……cho phép người dùng là những người X1 có khả năng thực thi được việc Y1.

Hơn nữa, nó còn cho phép người dung X2 có khả năng thực hiện việc Y2

…..

## Xác Định, Chuẩn hóa dữ liệu

Xét Ví dụ 2

Craw từ nguồn nào, sử dụng thư viện gì của môi trường nào (ví dụ Selenium python)

Tiếp đó, sử dụng công cụ, thư viện gì để xử lý việc gì (Ví dụ X path lấy hết bình luận, rồi xuất ra file CSV)

Thực hiện Tokenizer để làm chuẩn hóa dữ liệu (bao gồm những việc sau….)

Sau khi thực hiện các việc trên, dữ liệu được lưu thành từng cột và lưu vào file X

## Thiết kế và xây dựng mô hình

### Thiết kế

### Xây dựng mô hình

#### Đầu vào

#### Xử lý

Bước tiếp theo….sử dụng Matplotlib… để thực hiện…

Sentiment Analyis với thư viện NLTK….

#### Đầu ra:

Bảng phân loại cảm xúc của từng bình luận từ dữ liệu đầu vào

Biểu đồ thể hiện mức độ phần tram hoặc so sánh số lượng cảm xúc theo từng phân loại từ dữ liệu đầu vào. Ví dụ Tích cực chiếm 40% số lượng; Tiêu cực 30% số lượng

#### Xuất ra mô hình:

Lúc xây dựng mô hình, có thể mình có model được lưu trong một biến, từ biến này mình có thể gọi để truyền dữ liệu vào và chạy trực tiếp. Trường hợp này lập trình viên đang thao tác. Người dung không chạy bằng cách này được.

Từ biến này mình có thể lưu mô hình ra thành file. Để từ đó ứng dụng khác có thể gọi và chạy độc lập

## Thử nghiệm và đánh giá mô hình

### Làm thế nào để biết mô hình tốt hay chưa tốt?

Kiểm tra lại kết quả từ mô hình chạy ra với dữ liệu thực tế. Cần lập bảng so sánh.

Lấy ví dụ: Trong 100 bình luận mô hình chạy ra được Tích cực chiếm 40% số lượng; Tiêu cực 30% số lượng. Trong khi đó, với dữ liệu đã được gán nhãn, thì thực tế tích cực sẽ chiếm 50, tiêu cực 20%. Sự sai khác này biểu diễn như thế nào?

Trường hợp này, có thể sử dung công cụ Confusion matrix để thể hiện

### Thử và so sánh

Bởi vì mỗi một lần thay đổi các thông số đầu vào của mô hình, sẽ làm cho kết quả bị thay đổi, nó có thể là tốt hơn, có thể là xấu hơn.

Làm thế nào để biểu diễn các lần so sánh này

Lập một bảng ghi nhận các trường hợp chạy. Bảng này có thể gồm các cột:

Lần thử nghiệm, thử với đặc trưng gì (3 feature gồm, 4 feature gồm…), kết quả đạt được độ chính xác bao nhiêu? Mức độ nhầm lẫn bao nhiêu)

Từ bảng này cũng có thể vẽ biểu đồ, biểu đồ này sẽ thể hiện mức độ tốt xấu của mô hình trong các trường hợp chạy thử nghiệm.

### Kết luận

Từ bảng ở bước trên, từ biểu đồ ở bước trên, mình đưa ra nhận định là trường hợp nào tốt nhất, có thể sử dụng trong trường hợp nào. Có thể đưa ra lý giải vì sao trường hợp này có thể đạt được mức độ tốt như vậy. Lấy vị dụ trường hợp 4 features ở trên tốt vì lý do sau đây…

### Xuất mô hình tốt nhất

Trong số các trường hợp ở trên, trường hợp 4 feature cho kết quả tốt nhất phù hợp với ngữ cảnh sẽ được xuất ra và load vào hệ thống cho người sử dụng có thể thao tác được.

### Viết một đoạn code để có thể thực thi các trường hợp thử này

Gợi ý: Có thể xây dựng một file Excel/ csv các trường hợp thử, đưa ra tham số thử, load vào, cho chạy lặp lại cho đến khi hoàn thành các trường hợp này và lưu ngược lại kết quả đánh giá ra file.

Cuối cùng, chọn lấy một trường hợp gọi là phù hợp nhất chạy lại lần cuối và xuất mô hình ra.

Làm như vậy sẽ tự động hóa quá trình chạy và lưu mô hình tốt nhất phù hợp nhất ra file

## Xây dựng giao diện người sử dụng

### Xác định đầu vào và đầu ra cho người sử dụng

Thiế kế và xây dựng giao diện cho người lập trình có thể thử nghiệm được

Thiết kế và xây dựng giao diện người sử dụng có thể thao tác được.

Ai dung hệ thống đưa vào dữ liệu gì và nhận về dữ liệu gì?

Cần xác định dữ liệu đầu vào ví dụ như mỗi lần thao tác là một bản ghi. Hay mỗi lần thao tác là cả một tập bản ghi.

Như vậy, hệ thống sẽ cần phải được thiết kế để cho phép người dung nhập vào trường hợp 1 bản ghi (giao diện nhập) hay là trường hợp upload file csv chẳng hạn rồi trả về kết quả

### Xác định đầu ra

Đầu ra là dạng gì? Một kết quả kết luận, hay một tập hợp các kết luận, hay là một bảng biểu được biểu diễn trên giao diện trên Web chẳng hạn.

Từ đó đưa ra thiết kế giao diện

Để thực hiện việc này, tốt hơn hết là nên vẽ giao diện dự kiến (có thể sử dụng các công cụ vẽ mockup

## Vận hành thử nghiệm

Cần trình bày, module này đã được tích hợp vào đâu, hệ thống nào và đã cho kết quả ra sao.

# MỘT TRƯỜNG hợp Phân tích cảm xúc bình luận sản phẩm

## Mô tả chung

Xây dựng một kịch bản, bài toán ứng dụng trong thực tế.

Cần nhấn mạnh giải quyết được một bài toán cụ thể, một tình huống cụ thể trong thực tế mà người làm công việc đó sẽ thực hiện.

Khi chưa có AI thì hoạt động này đang triển khai như thế nào, cụ thể nêu quy trình (bao gồm các bước hoặc ngữ cảnh mà bước đó sẽ thực hiện), từ đó đi sâu vào mô tả chi tiết hoạt động mà mình muốn nghiên cứu và triển khai.

Hình thức của đề tài có thể là một sản phẩm hoàn chỉnh đã được thực thi và triển khai thử nghiệm, nhưng cũng có thể chấp nhận trường hợp chưa lập trình xong nhưng cần được bù đắp và giải thích bằng thực tế tại sao lại chưa triển khai được. Trường hợp này thường là có ý tưởng tốt, nhưng việc lấy dữ liệu có thể triển khai gặp khó khăn, hoặc mô hình cần những điểm mới mà mình đã chỉ ra trong phần thiết kế một cách cụ thể tuy nhiên chưa phát triển xong. Điều này, cần phải được nhấn mạnh trong phần tóm tắt của đề tài nhóm.

## Xác định bài toán

### Tên đề tài

Ví dụ Xét bài toán phân tích cảm xúc từ các bình luận trên hệ thống web thường mại điện tử

Ví dụ Bài viết trên Fanpage về dịch vụ XYZ.

### Mục tiêu cùa đề tài

*Đề tài này nhắm đến mục đích gì:*

Xây dựng được một mô hình ….để thực hiện ……cho phép người dùng là những người X1 có khả năng thực thi được việc Y1.

Hơn nữa, nó còn cho phép người dung X2 có khả năng thực hiện việc Y2

…..

## Xác Định, Chuẩn hóa dữ liệu

Xét Ví dụ 2

Craw từ nguồn nào, sử dụng thư viện gì của môi trường nào (ví dụ Selenium python)

Tiếp đó, sử dụng công cụ, thư viện gì để xử lý việc gì (Ví dụ X path lấy hết bình luận, rồi xuất ra file CSV)

A screenshot of a phone

Description automatically generated

Thực hiện Tokenizer để làm chuẩn hóa dữ liệu (bao gồm những việc sau….)

Sau khi thực hiện các việc trên, dữ liệu được lưu thành từng cột và lưu vào file X

## Thiết kế và xây dựng mô hình

### Thiết kế

Thiết kế về quy trình xây dựng và đánh giá

A screenshot of a computer

Description automatically generated

### Xây dựng mô hình

#### Đầu vào

#### Xử lý

Bước tiếp theo….sử dụng Matplotlib… để thực hiện…

Sentiment Analyis với thư viện NLTK….

#### Đầu ra:

Bảng phân loại cảm xúc của từng bình luận từ dữ liệu đầu vào

Biểu đồ thể hiện mức độ phần tram hoặc so sánh số lượng cảm xúc theo từng phân loại từ dữ liệu đầu vào. Ví dụ Tích cực chiếm 40% số lượng; Tiêu cực 30% số lượng

#### Xuất ra mô hình:

Lúc xây dựng mô hình, có thể mình có model được lưu trong một biến, từ biến này mình có thể gọi để truyền dữ liệu vào và chạy trực tiếp. Trường hợp này lập trình viên đang thao tác. Người dung không chạy bằng cách này được.

Từ biến này mình có thể lưu mô hình ra thành file. Để từ đó ứng dụng khác có thể gọi và chạy độc lập

## Thử nghiệm và đánh giá mô hình

### Làm thế nào để biết mô hình tốt hay chưa tốt?

Kiểm tra lại kết quả từ mô hình chạy ra với dữ liệu thực tế. Cần lập bảng so sánh.

Lấy ví dụ: Trong 100 bình luận mô hình chạy ra được Tích cực chiếm 40% số lượng; Tiêu cực 30% số lượng. Trong khi đó, với dữ liệu đã được gán nhãn, thì thực tế tích cực sẽ chiếm 50, tiêu cực 20%. Sự sai khác này biểu diễn như thế nào?

Trường hợp này, có thể sử dung công cụ Confusion matrix để thể hiện

### Thử và so sánh

Bởi vì mỗi một lần thay đổi các thông số đầu vào của mô hình, sẽ làm cho kết quả bị thay đổi, nó có thể là tốt hơn, có thể là xấu hơn.

Làm thế nào để biểu diễn các lần so sánh này

Lập một bảng ghi nhận các trường hợp chạy. Bảng này có thể gồm các cột:

Lần thử nghiệm, thử với đặc trưng gì (3 feature gồm, 4 feature gồm…), kết quả đạt được độ chính xác bao nhiêu? Mức độ nhầm lẫn bao nhiêu)

Từ bảng này cũng có thể vẽ biểu đồ, biểu đồ này sẽ thể hiện mức độ tốt xấu của mô hình trong các trường hợp chạy thử nghiệm.

Ví dụ

Đối với mô hình CNN

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| learning rate | filters | kernel\_size | batch size | epoch | Độ chính xác trung bình | Độ lệch chuẩn |
| 0.001 | 32 | 32 | 128 | 10 | 0.734554 | 0.0544791 |
| 15 | 0.763827 | 0.015081 |
| 20 | 0.763827 | 0.015081 |
| 256 | 10 | 0.680618 | 0.0346901 |
| 15 | 0.660937 | 0.059854 |
| 20 | 0.750997 | 0.01185 |
| 16 | 128 | 10 | 0.714748 | 0.0736507 |
| 15 | 0.77728 | 0.00767662 |
| 20 | 0.775536 | 0.0154148 |
| 256 | 10 | 0.669905 | 0.0447696 |
| 15 | 0.764574 | 0.0171329 |
| 20 | 0.775411 | 0.00600238 |

Bảng 3.1: Tổng hợp độ chính xác trung bình và độ lệch chuẩn với các tham số mô hình CNN

Đối với mô hình biLSTM

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| learning rate | lstm\_units | batch size | epoch | Độ chính xác trung bình | Độ lệch chuẩn |
| 0.001 | 32 | 128 | 10 | 0.787369 | 0.00169883 |
| 15 | 0.785252 | 0.000932152 |
| 20 | 0.785999 | 0.00459032 |
| 256 | 10 | 0.785127 | 0.00132998 |
| 15 | 0.792352 | 0.000635154 |
| 20 | 0.790359 | 0.00721397 |
| 16 | 128 | 10 | 0.791106 | 0.00139823 |
| 15 | 0.797583 | 0.00366565 |
| 20 | 0.787494 | 0.00493248 |
| 256 | 10 | 0.787743 | 0.00230359 |
| 15 | 0.788366 | 0.00438281 |
| 20 | 0.790732 | 0.00461728 |

Bảng 3.2: Tổng hợp độ chính xác trung bình và độ lệch chuẩn với các tham số mô hình BiLSTM

Phương pháp đo lường

Trong nghiên cứu này, để đo lường kết quả cũng như so sánh các mô hình nghiên cứu sẽ xem xét các chỉ số: **Accuracy**, **Precision**, **Recall**, **F1-score**, và **Loss.** Trong đó,

* Accuracy (Độ chính xác): Tỷ lệ giữa số lượng dự đoán đúng và tổng số dự đoán. Đây là chỉ số tổng quát phản ánh mức độ chính xác của mô hình. Nếu độ chính xác (Accuracy) là 85%, điều này có nghĩa là trong tổng số dự đoán của mô hình, 85% trong số đó là đúng.
* Precision (Độ chính xác dương): Tỷ lệ giữa số lượng dự đoán đúng thuộc lớp dương và tổng số dự đoán thuộc lớp dương. Precision cho thấy mức độ chính xác khi mô hình dự đoán một đối tượng là thuộc lớp dương. Nếu Precision là 80%, điều này có nghĩa là trong số tất cả các dự đoán mà mô hình phân loại là dương, 80% thực sự đúng.
* Recall (Độ nhạy): Tỷ lệ giữa số lượng dự đoán đúng thuộc lớp dương và tổng số thực tế thuộc lớp dương. Recall đo lường khả năng của mô hình trong việc phát hiện các trường hợp dương. Nếu Recall là 75%, điều này có nghĩa là trong số tất cả các trường hợp thực tế thuộc lớp dương, mô hình chỉ phát hiện được 75%.
* F1-score: Là trung bình điều hòa giữa Precision và Recall. Chỉ số này đặc biệt hữu ích khi có sự mất cân bằng giữa các lớp, giúp cân bằng giữa độ chính xác và độ nhạy. Nếu F1-score là 77%, điều này cho thấy mô hình có sự cân bằng tốt giữa độ chính xác dương và độ nhạy.
* Loss (Mất mát): Đây là chỉ số đo lường sự khác biệt giữa dự đoán của mô hình và giá trị thực tế. Loss thường được sử dụng để tối ưu hóa mô hình trong quá trình huấn luyện. Nếu Loss là 0.35, điều này cho thấy mức độ sai lệch giữa dự đoán của mô hình và giá trị thực tế. Giá trị loss thấp hơn thường chỉ ra rằng mô hình hoạt động hiệu quả.

### Kết quả thử nghiệm với tập dữ liệu kiểm tra và dữ liệu mới

Kết quả các trường hợp thử nghiệm được tập hợp ở bảng sau

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | | Tập kiểm tra | | | Tập dữ liệu mới | | |
| CNN | BiLSTM | CNN-  BiLSTM | CNN | BiLSTM | CNN-  BiLSTM |
| Accuracy | | 0.81 | 0.81 | 0.82 | 0.84 | 0.85 | 0.87 |
| Precision | Tiêu cực | 0.83 | 0.85 | 0.82 | 0.82 | 0.87 | 0.90 |
| Trung lập | 0.75 | 0.76 | 0.78 | 0.79 | 0.84 | 0.86 |
| Tích cực | 0.84 | 0.84 | 0.85 | 0.88 | 0.89 | 0.88 |
| Recall | Tiêu cực | 0.80 | 0.82 | 0.85 | 0.84 | 0.91 | 0.86 |
| Trung lập | 0.75 | 0.76 | 0.76 | 0.80 | 0.80 | 0.81 |
| Tích cực | 0.85 | 0.86 | 0.86 | 0.87 | 0.87 | 0.92 |
| F1-score | Tiêu cực | 0.81 | 0.83 | 0.83 | 0.83 | 0.89 | 0.88 |
| Trung lập | 0.75 | 0.76 | 0.77 | 0.79 | 0.79 | 0.81 |
| Tích cực | 0.85 | 0.85 | 0.85 | 0.88 | 0.88 | 0.90 |

Bảng 4.1: Mô tả kết quả trên tập kiểm tra - tập dữ liệu mới

Bảng 4.1 so sánh kết quả thử nghiệm trên ba mô hình CNN, biLSTM và CNN - BiLSTM với tập dữ liệu kiểm tra trích xuất từ tập dữ liệu kiểm thử và tập dữ liệu mới gồm 300 câu là những đánh giá sản phẩm chưa từng được đưa vào mô hình trước đó. Bảng dữ liệu so sánh dựa trên bốn tham số là accuracy (độ chính xác), percision, recall và f1-score với ba lớp phân loại: “Tiêu cực”, “Trung lập”, “Tích cực”.

Xét về accuracy (độ chính xác), ở tập dữ liệu kiểm tra hai mô hình CNN – BiLSTM mang lại kết quả tốt nhất là 0.82 theo sau là BiLSTM và CNN với 0.81 dường như không có sự chênh lệch nhiều; ở tập dữ liệu mới mô hình CNN – BiLSTM cho ra kết quả cao nhất với 0.87 tiếp đó là BiLSTM với 0.85. Điều này cho thấy trong cả hai tập kiểm tra lẫn dữ liệu mới đều mô hình CNN – BiLSTM mang lại hiệu suất tốt nhất trong ba mô hình CNN, BiLSTM và CNN – BiLSTM.

Xét về precision, trên tập kiểm tra ta dễ dàng nhận thấy mô hình CNN - BiLSTM đem đến hiệu suất cao nhất 0.85 việc xác định “Tích cực” và kém nhất hơn ở “Trung lập” 0.78 nhưng vẫn tốt hơn kết quả CNN ở lớp “Trung lập” với 0.75 nhưng lại kém nhất trong ba mô hình với lớp “Tiêu cực”. Ngoài CNN - BiLSTM thì BiLSTM cũng đem lại kết quả tốt cả ba lớp đặc biệt ở lớp “Tiêu cực” 0.85. Trên tập dữ liệu mới, ta thấy mô hình CNN – biLSTM có lại hiệu suất cao nhất trong việc phân loại lớp “Tiêu cực” đến với 0.9 và đứng phía sau là BiLSTM với 0.87. Các kết quả các của hai mô hình này đều tương đối đồng đều chỉ có CNN lại có sự chêch lệch khá đáng kể giữa các lớp khi đạt 0.82 ở “Tiêu cực”, 0.88 ở “Tích cực” nhưng lại chỉ có 0.79 ở “Trung lập”. Từ đó ta có thể thấy mô hình CNN-BiLSTM không chỉ đạt hiệu suất tốt mà còn mang lại sự đồng đều trong việc phân loại các lớp cảm xúc, đặc biệt trên tập dữ liệu mới. Mô hình BiLSTM cũng cho thấy khả năng xử lý tốt cả ba lớp, tuy nhiên, vẫn kém hơn CNN-BiLSTM một chút ở lớp “Tiêu cực”. Trong khi đó, CNN lại có sự chênh lệch rõ rệt giữa các lớp, với kết quả thấp nhất ở lớp “Trung lập”. Điều này cho thấy CNN có thể thiếu khả năng xử lý thông tin tuần tự và biểu diễn ngữ nghĩa phức tạp, dẫn đến hiệu suất không đồng đều.

Xét về Recall, trên tập dữ liệu kiểm tra mô hình BiLSTM và CNN - BiLSTM cho thấy khả năng bỏ sót của hai mô hình này khá tương đồng nhau đều là 0.86 ở lớp “Tích cực” , 0.76 ở lớp “Trung lập” nhưng CNN – BiLSTM lại cho kết quả tốt hơn ở lớp “Tiêu cực” với 0.85. Trong khi đó, lần này kết quả của CNN đều thấp hơn hai mô hình còn lại ở tất cả các lớp nhưng khoảng cách chỉ là rất nhỏ. Ở tập dữ liệu mới CNN – BiLSTM cũng đem đến kết quả xuất sắc khi đứng đầu ở các lớp “Tích cực” 0.92 và “Trung lập” 0.81 và chỉ kém mô hình BiLSTM ở lớp “Tiêu cực”. Ở tập dữ liệu này BiLSTM đem đến kết quả 0.91 ở lớp “Tích cực” và tương đồng với CNN ở hai lớp “Trung lập” và “Tiêu cực”. Ta thấy sự cách biệt giữa các mô hình đặc biệt ở lớp “Trung lập” là không đáng kể chỉ cách 0.01. Những điều này cho thấy mô hình CNN-BiLSTM vượt trội trong việc nhận diện các lớp cảm xúc, BiLSTM cũng thể hiện hiệu suất rất tốt, đặc biệt ở lớp “Tiêu cực” trên tập dữ liệu mới, nhưng vẫn kém đồng đều hơn CNN-BiLSTM. Còn CNN mặc dù đạt kết quả thấp hơn nhưng khoảng cách so với hai mô hình còn lại không quá lớn, cho thấy nó vẫn có khả năng cạnh tranh.

Xét về F1-score, mô hình CNN cho ra kết quả thấp nhất ở tập kiểm tra với lớp “Trung lập” với 0.75 và tương tự với tập dữ liệu mới ở lớp này cùng với mô hình BiLSTM. Trong khi đó, ở tập kiểm tra thì CNN – BiLSTM và BiLSTM lại đế đến kết quả cao nhất ở lớp “Tiêu cực” là 0.83 nhưng ở tập dữ liệu mới thì mô hình BiLSTM lại mang đến kết quả tốt hơn ở lớp này với 0.89 theo sau là 0.88 của mô hình CNN – biLSTM. Dựa trên bảng số liệu ta cũng có thể thấy mô hình CNN – biLSTM đem đến các giá trị F1-score ở các lớp đều khá tốt cho thấy phần vượt trội hơn so với hai mô hình còn lại tiếp theo phía sau là BiLSTM còn cuối cùng vẫn là CNN. Ở hai tập dữ liệu kiểm tra và dữ liệu mới kết quả của CNN và BiLSTM đều khá tương đồng ở lớp “Tích cực” và “Trung lập” nhưng ở lớp “Tích cực” BiLSTM lại có phần vượt trội hơn có khi vượt qua cả CNN – BiLSTM. Tuy nhiên, CNN-BiLSTM tiếp tục khẳng định ưu thế vượt trội của mình với kết quả F1-score đồng đều và cao nhất ở hầu hết các lớp.

### Ví dụ Kết quả đào tạo huấn luyện

#### Mô hình CNN

A graph of a line and a line

Description automatically generated with medium confidence

Hình 4.1: Kết quả đào tạo mô hình CNN

Hình 4.1 mô tả kết quả huấn luyện của mô hình CNN qua 15 epoch. Kết quả cho thấy mô hình huấn luyện đã có sự cải thiện đáng kể về hiệu suất. Từ một khả năng dự đoán khá thấp ở epoch đầu tiên, mô hình đã trải qua quá trình học và phát triển, đạt đến mức độ ổn định và chính xác đáng kể. Cụ thể, ở tập huấn luyện giá trị của hàm mất mát – loss (đường màu đỏ) luôn giữ nguyên xu hướng giảm cụ thể giá trị này đã giảm mạnh từ 1.0602 ở epoch số 1 xuống chỉ còn 0.2241 ở epoch số 15. Đồng thời, độ chính xác – accuracy (đường màu xanh) của mô hình cũng có xu hướng tăng lên từ 0.4171 ở epoch số 1, nó đã tăng lên đến 0.9067 ở epoch số 15. Đây dấu hiệu tích cực, cho thấy mô hình đang học hiệu quả và hiệu suất của nó đang được cải thiện trong thời gian huấn luyện. Điều này chỉ ra rằng mô hình không chỉ có hiệu suất tốt trên tập huấn luyện có sự cải thiện hiệu suất tương đối ổn định qua từng epoch có khả năng học tốt từ dữ liệu huấn luyện. Quá trình giảm loss song hành cùng việc tăng accuracy phản ánh sự tối ưu hóa hiệu quả của các tham số trong mô hình.

### Kết luận

Từ bảng ở bước trên, từ biểu đồ ở bước trên, mình đưa ra nhận định là trường hợp nào tốt nhất, có thể sử dụng trong trường hợp nào. Có thể đưa ra lý giải vì sao trường hợp này có thể đạt được mức độ tốt như vậy. Lấy vị dụ trường hợp 4 features ở trên tốt vì lý do sau đây…

### Xuất mô hình tốt nhất

Trong số các trường hợp ở trên, trường hợp 4 feature cho kết quả tốt nhất phù hợp với ngữ cảnh sẽ được xuất ra và load vào hệ thống cho người sử dụng có thể thao tác được.

### Viết một đoạn code để có thể thực thi các trường hợp thử này

Gợi ý: Có thể xây dựng một file Excel/ csv các trường hợp thử, đưa ra tham số thử, load vào, cho chạy lặp lại cho đến khi hoàn thành các trường hợp này và lưu ngược lại kết quả đánh giá ra file.

Cuối cùng, chọn lấy một trường hợp gọi là phù hợp nhất chạy lại lần cuối và xuất mô hình ra.

Làm như vậy sẽ tự động hóa quá trình chạy và lưu mô hình tốt nhất phù hợp nhất ra file

## Xây dựng giao diện người sử dụng

### Đề xuất kiến trúc hệ thống

A white rectangular object with black text

Description automatically generated

Đề xuất quy trình tích hợp khi có sự tham gia của module mới

Từ người dung quy trình sẽ như thế nào? Ví dụ

A black and white screen shot of a computer screen

Description automatically generated

Từ quản trị sẽ thao tác như thế nào? Ví dụ

A group of white rectangular objects

Description automatically generated

### Xác định đầu vào và đầu ra cho người sử dụng

Thiế kế và xây dựng giao diện cho người lập trình có thể thử nghiệm được

Thiết kế và xây dựng giao diện người sử dụng có thể thao tác được.

Ai dung hệ thống đưa vào dữ liệu gì và nhận về dữ liệu gì?

Cần xác định dữ liệu đầu vào ví dụ như mỗi lần thao tác là một bản ghi. Hay mỗi lần thao tác là cả một tập bản ghi.

Như vậy, hệ thống sẽ cần phải được thiết kế để cho phép người dung nhập vào trường hợp 1 bản ghi (giao diện nhập) hay là trường hợp upload file csv chẳng hạn rồi trả về kết quả

### Xác định đầu ra

Đầu ra là dạng gì? Một kết quả kết luận, hay một tập hợp các kết luận, hay là một bảng biểu được biểu diễn trên giao diện trên Web chẳng hạn.

Từ đó đưa ra thiết kế giao diện

Để thực hiện việc này, tốt hơn hết là nên vẽ giao diện dự kiến (có thể sử dụng các công cụ vẽ mockup

### Ví dụ giao diện người dung

Trong giao diện này, bình luận của AI được hiện lên sau khi đã gọi vào module phân loại

A screenshot of a phone

Description automatically generated

### Quản trị có thể thấy như thế nào

A screenshot of a computer

Description automatically generated

## Vận hành thử nghiệm

Cần trình bày, module này đã được tích hợp vào đâu, hệ thống nào và đã cho kết quả ra sao.