

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN



**BÁO CÁO ĐỒ ÁN CUỐI KỲ
MÔN: XỬ LÝ NGÔN NGỮ TỰ NHIÊN**

**ĐỀ TÀI: PHÂN TÍCH CẢM XÚC VỀ CÁC ĐÁNH GIÁ SẢN PHẨM
BẰNG TIẾNG ANH TRÊN AMAZON**

Giảng viên hướng dẫn: TS. Nguyễn Trọng Chính

Lớp: CS221.P21

Sinh viên thực hiện:

STT	Họ và tên	MSSV
1	Lương Anh Huy	22520550
2	Hồng Khải Nguyên	22520967
3	Lê Bá Quốc	23521303

MỤC LỤC

CHƯƠNG 1: GIỚI THIỆU BÀI TOÁN	1
1.1 Lý do lựa chọn đề tài	1
1.2 Phát biểu bài toán	2
CHƯƠNG 2: BỘ NGỮ LIỆU	3
2.1 Thu thập ngữ liệu	3
2.2 Mô tả nhãn của ngữ liệu	3
2.3 Thống kê ngữ liệu	5
2.4 Kiểm tra và giải thích các kết quả gán nhãn	8
CHƯƠNG 3: PHƯƠNG PHÁP	35
3.1 Kỹ thuật trích xuất đặc trưng	35
3.1.1 Term Frequency – Inverse Document Frequency (TF-IDF)	35
3.2 Các mô hình sử dụng.....	37
3.2.1 Multinomial Naïve Bayes	37
3.2.2 Logistic Regression	39
3.2.3 Linear Support Vector Machine - LinearSVM.....	42
3.3 Các độ đo đánh giá mô hình	44
3.3.1 Accuracy	44
3.3.2 Precision	44
3.3.3 Recall	44
3.3.4 F1 – Score	45
3.3.5 Weighted F1 – Score	45
CHƯƠNG 4: THỰC NGHIỆM.....	46
4.1 Tiền xử lý dữ liệu	46
4.1.1 Xóa các mẫu bị trùng lặp và các mẫu không có đánh giá	46
4.1.2 Trích xuất và loại bỏ	46
4.1.3 Chuẩn hóa và ánh xạ.....	48
4.1.4 Sửa lỗi chính tả	51
4.2 Huấn luyện và kiểm thử các mô hình	52
4.2.1 MNB + TF – IDF.....	52
4.2.2 LR + TF – IDF	53
4.2.3 SVM + TF – IDF	53

CHƯƠNG 5: NHẬN XÉT VÀ PHÂN TÍCH KẾT QUẢ	55
5.1 Multinomial Naïve Bayes	55
5.2 Logistic Regression	58
5.3 SVM	62
CHƯƠNG 6: KẾT LUẬN	66
TÀI LIỆU THAM KHẢO.....	68

MỤC LỤC HÌNH ẢNH

Bảng 1.1 Ví dụ minh họa cho phát biểu bài toán.....	2
Bảng 2.1 Bảng mô tả nhãn của dữ liệu	4
Bảng 2.2 Thống kê tổng quan về bộ ngữ liệu	5
Bảng 2.3 Kiểm tra và giải thích gán nhãn của 72 mẫu	8
Bảng 3.1 Ví dụ về cách tính của kỹ thuật TF - IDF	36
Bảng 4.1 Giá trị tham số tối ưu của MNB và TF-IDF.....	52
Bảng 4.2 Classification Report của mô hình MNB.....	52
Bảng 4.3 Giá trị tham số tối ưu của mô hình LR và TF-IDF	53
Bảng 4.4 Classification Report của mô hình LR	53
Bảng 4.5 Giá trị tham số tối ưu của mô hình SVM	53
Bảng 4.6 Classification Report của mô hình SVM	54
Bảng 4.7 Tóm tắt kết quả của ba mô hình	54
Bảng 5.1 Giá trị về các hệ số đặc trưng của từng lớp trong mô hình MNB.....	55
Bảng 5.2 Một số trường hợp MNB dự đoán nhãn khác với nhãn thực tế.....	56
Bảng 5.3 Giá trị về các hệ số đặc trưng của từng lớp trong mô hình LR	59
Bảng 5.4 Một số trường hợp LR dự đoán nhãn khác với nhãn thực tế	60
Bảng 5.5 Giá trị về các hệ số đặc trưng của từng lớp trong mô hình SVM	62
Bảng 5.6 Một số trường hợp SVM dự đoán nhãn khác với nhãn thực tế	63

MỤC LỤC HÌNH ẢNH

Hình 2.1 Một số ngữ liệu trong tập dữ liệu	3
Hình 2.2 Phân bố nhãn trên từng lớp của bộ dữ liệu	6
Hình 2.3 Top 20 từ xuất hiện phổ biến trong tập dữ liệu	7
Hình 3.1 Ví dụ thực tế về giá trị TF-IDF trên một mẫu trong bộ ngữ liệu	36
Hình 3.2 Mã giả của mô hình Multinomial Naïve Bayes	38
Hình 3.3 Mã giả của mô hình Logistic Regression với bài toán phân loại nhị phân	40
Hình 3.4 Mã giả của mô hình Logistic Regression cho bài toán phân loại đa lớp	41
Hình 3.5 Mã giả của mô hình LinearSVM	43

CHƯƠNG 1: GIỚI THIỆU BÀI TOÁN

1.1 Lý do lựa chọn đề tài

Trong thời đại công nghệ phát triển như hiện nay, các giao dịch, mua bán, trao đổi sản phẩm, hàng hóa không cần phải thực hiện một cách trực tiếp mà dần dần chuyển sang hình thức trực tuyến, tiết kiệm chi phí đi lại và thời gian của nhiều người. Các nền tảng thương mại điện tử dần trở nên chiếm ưu thế hơn so với các cửa hàng truyền thống bởi tính tiện lợi, nhanh chóng và gọn nhẹ, ngoài ra trên các nền tảng thường có những ưu đãi, giảm giá rẻ hơn so với các sản phẩm được bán trực tiếp. Ở Việt Nam thì Shopee đang là mạnh nhất, sau đó là Tiki, Lazada, Amazon,... Với mỗi sản phẩm được bán, sẽ có những bình luận, đánh giá về chất lượng sản phẩm từ người mua trước đó, những người đang có nhu cầu có thể xem qua để cân nhắc chọn lựa và các doanh nghiệp có thể xem qua để hiểu hơn về chất lượng sản phẩm của mình trong tay khách hàng. Những bình luận và đánh giá đó là nguồn dữ liệu phong phú và giá trị cho cả người tiêu dùng lẫn các doanh nghiệp.

Phân tích cảm xúc là một trong những bài toán điển hình và trọng tâm trong lĩnh vực Xử lý ngôn ngữ tự nhiên, với mục tiêu là xác định và phân loại cảm xúc được thể hiện trong các văn bản. Cụ thể hơn, bài toán này giúp cho các hệ thống máy tính có khả năng tự động gán nhãn cho câu văn hoặc đoạn văn bản theo các mức độ cảm xúc. Việc phân tích cảm xúc tự động và nhanh chóng mang lại nhiều ứng dụng thực tế quan trọng, có thể đánh giá phản ứng và nhận xét của người dùng đối với các sản phẩm vừa ra mắt hoặc các chiến dịch truyền thông, hoặc phát hiện, theo dõi các xu hướng dư luận trên mạng xã hội về một chủ đề, đề tài cụ thể. Ngoài ra còn có thể hỗ trợ ra các quyết định trong các lĩnh vực như tài chính, marketing, chăm sóc khách hàng,...

Trong đồ án này, nhóm thực hiện phân tích cảm xúc trên dữ liệu là các đánh giá sản phẩm giày dép bằng tiếng Anh trên nền tảng thương mại Amazon, mỗi đánh giá sẽ đi kèm với số sao từ 1 đến 5 tương ứng thể hiện mức độ từ rất không hài lòng đến rất hài lòng của khách hàng. Để phù hợp với bài toán phân tích cảm xúc, nhóm phân loại 5 loại đánh giá thành 3 cảm xúc chính và tổng quát nhất là tích cực (positive) gồm các đánh giá từ 4 đến 5 sao, trung tính (neutral) gồm các đánh giá 3 sao và tiêu cực (negative) gồm các đánh giá từ 1 đến 2 sao. Bên cạnh đó nhóm sẽ sử dụng một mô hình học sâu

làm phương pháp chính và so sánh hiệu năng và kết quả với một mô hình học máy kết hợp với kỹ thuật trích xuất đặc trưng trong xử lý ngôn ngữ tự nhiên

1.2 Phát biểu bài toán

Trong phạm vi đồ án, bài toán phân tích cảm xúc được định nghĩa như sau

- Input: Một đoạn bình luận hoặc đánh giá của khách hàng về sản phẩm bằng tiếng Anh
- Output: Một nhãn cảm xúc tương ứng với bình luận đầu vào, được phân loại thành một trong ba lớp là tích cực, tiêu cực hoặc trung tính

Ví dụ minh họa

Bảng 1.1 Ví dụ minh họa cho phát biểu bài toán

Bình luận đầu vào	Nhãn cảm xúc trả về
Love this bag! Perfect size for my wallet, cellphone, keys and sunglasses. The lining is striped so I can find things easier than a solid dark color.	Positive
Two left found in the box ..	Neutral
bad quality the color went off after 2 weeks	Negative

CHƯƠNG 2: BỘ NGỮ LIỆU

2.1 Thu thập ngữ liệu

Bộ dữ liệu mà nhóm sử dụng trong đề tài được thu thập từ nền tảng HuggingFace, một kho lưu trữ uy tín với hàng nghìn bộ dữ liệu và mã nguồn mở phục vụ cho các tác vụ liên quan đến máy học nói chung và xử lý ngôn ngữ tự nhiên nói riêng. Cụ thể, nhóm sử dụng bộ dữ liệu Amazon Shoe Reviews, được cung cấp tại đường dẫn: <https://huggingface.co/datasets/juliensimon/amazon-shoe-reviews>

Bộ dữ liệu được tải trực tiếp bằng cách sử dụng thư viện datasets của HuggingFace trong môi trường Python. Tổng số mẫu của bộ dữ liệu là 100000 trong đó 90000 mẫu được sử dụng cho huấn luyện và 10000 mẫu được sử dụng cho kiểm tra

A	labels	text
1	3	Good shoe for office work. They will scuff very easy so be aware.
2	1	I have had the Patricia II wedge in black for about 1 year & wore them regularly in season. When I saw the Patricia II again, I was surprised at how well they still looked.
3	1	Width not right and size too small if width had been just little wider and ordered size larger would have been good.
4	0	I received these shoes and they weren't the same as the picture described them, they were a different color. When I tried them on, they were very uncomfortable.
5	2	They began to split along the mesh material after a month but loved the shoe and the feel of it.
6	4	Excellent shoes , very comfortable and litghweight !
7	0	Usually love Ethnies product. In this case the raised arch area of one shoe is too far back and raised to the point of being uncomfortable.
8	1	Could not get my foot into the shoe. Was disappointed and returned them.
9	3	Everything about the boot is great.
10	2	Nice looking shoe, okay for short-term wear. Much narrower than other size 11's that I have - tight fit!
11	2	The shoe was the EXACT color I was looking for but... You get exactly what you pay for. Poorly made shoe. Fit con
12	2	I loved how comfortable these shoes felt when I put them on. However, after a while, my big toe rubs against the toe box.
13	4	Slippers fit great. Nice and cozy.
14	2	These shoes are cute as several people have said, however, they are cut low and do not lace all the way up to your ankles.
15	1	Pros: They are very comfortable. I wear them all day long, every day with no complaints. They are also well ventilated.
16	0	These look great, the lenses are nice color and mirrored, and the matte finish is really nice. The good stops there.
17		

Hình 2.1 Một số ngữ liệu trong tập dữ liệu

Các đánh giá từ 1 đến 5 sao ở cột label được biểu diễn lần lượt dưới dạng các số nguyên từ 0 đến. Như đã đề cập ở phần lý do lựa chọn đề tài, nhóm thực hiện chuyển đổi các đánh giá từ 1 đến 2 sao thành nhãn Negative, 3 sao thành nhãn Neutral và 4 đến 5 sao thành nhãn Positive

2.2 Mô tả nhãn của ngữ liệu

Sau khi thực hiện chuyển đổi các đánh giá trên thang đo điểm số sao thành các nhãn cảm xúc, bộ dữ liệu gồm 3 nhãn là Positive, Neutral và Negative đại diện cho 3 trạng thái cảm xúc của khách hàng đối với sản phẩm:

- Positive: Biểu thị cảm xúc hài lòng hoặc đánh giá tốt về sản phẩm. Các bình luận thường chứa các từ ngữ mang tính khen ngợi như love, fit, comfortable, perfect,...
- Neutral: Thể hiện cảm xúc trung lập, không rõ ràng. Người dùng có thể vừa khen vừa chê trong cùng một đánh giá, hoặc chỉ đơn thuần cung cấp thông tin như chất liệu, kích cỡ,... mà không thể hiện cảm xúc rõ rệt.
- Negative: Biểu thị sự không hài lòng hoặc đánh giá kém về sản phẩm. Các bình luận thường xuất hiện các tính từ như hurt, disappoint, scratchy,...

Về mặt đặc trưng của ngôn ngữ, các đánh giá ở mức 1 hoặc 5 sao thường thể hiện cảm xúc rõ ràng, dễ nhận diện bởi các tính từ mạnh. Ngược lại ở mức 2 hoặc 4 sao, sắc thái trong đánh giá nhẹ nhàng hơn, ít sử dụng các tính từ mạnh nhưng vẫn thể hiện được thái độ tích cực hoặc tiêu cực. Cuối cùng, với mức 3 sao, đa phần khó có thể phân biệt thái độ rõ ràng của người tiêu dùng, họ có thể khen ở về này nhưng lại phê bình ở về kia, hoặc chỉ đơn giản là cung cấp thông tin một cách thông thường. Một số ví dụ minh họa rõ ràng cho các nhãn

Bảng 2.1 Bảng mô tả nhãn của dữ liệu

Đánh giá của khách hàng	Nhãn cảm xúc	Nhận xét của nhóm
Good shoe for office work. They will scuff very easy so be aware.	Positive	Đây là đánh giá rõ ràng khi khách hàng thể hiện thái độ yêu thích bởi sự phù hợp của giày cho công việc văn phòng
Good sandals for around the pool but not a pair I'd wear the whole day through. The footbed, as the listing called it, is very cushy and the rubber feels like memory foam. Recommended.	Positive	Mặc dù vừa có khen ngợi và hơi phê bình trong đánh giá nhưng phần khen ngợi vẫn chiếm nhiều hơn so với phê bình, người mua vẫn đề xuất cho các khách hàng khác mua

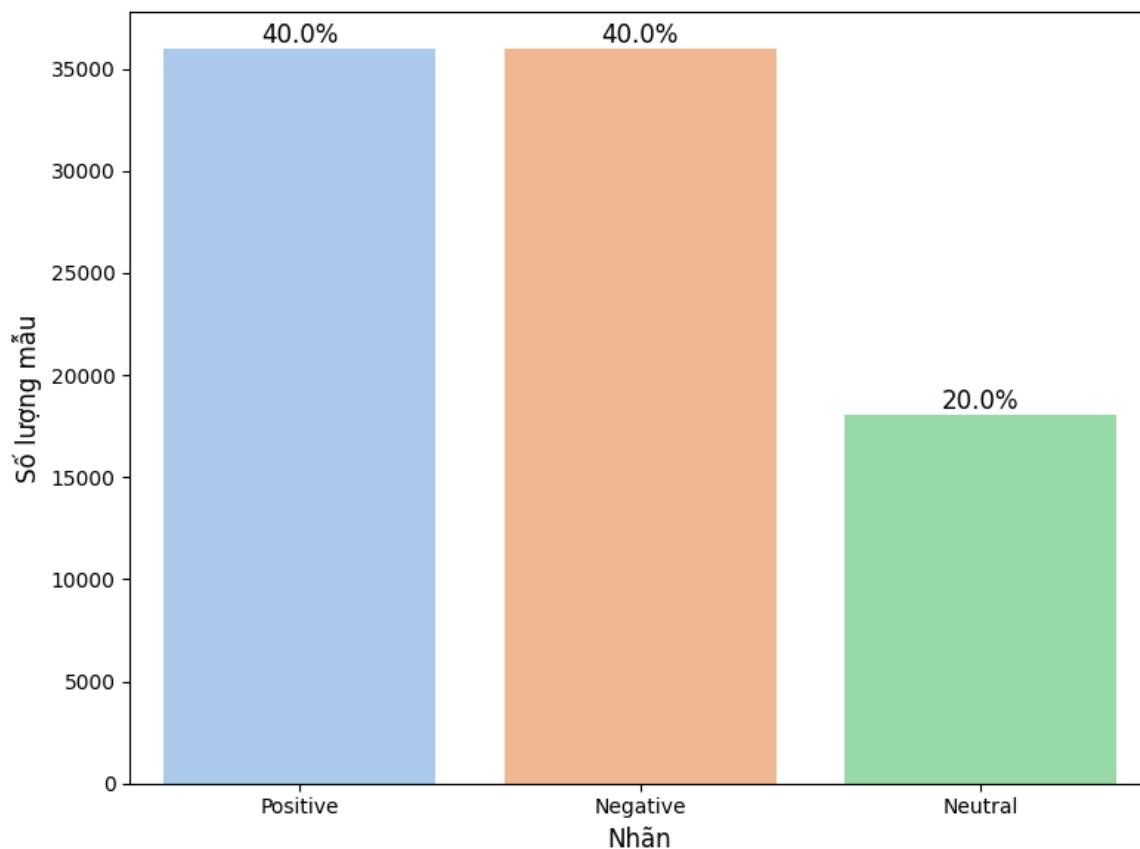
Could not get my foot into the shoe. Was disappointed and returned them.	Negative	Đây là đánh giá rõ ràng về sự thất vọng của khách hàng đối với đôi giày cũ thê
Were comfortable the 1st time, but seem to get tighter everytime I wore them 😞 recd many compliments on them but can no longer wear them	Negative	Sự yêu thích của khách hàng dần dần chuyển thành phê bình và không hài lòng về chất lượng sản phẩm
They began to split alone the mesh material after a month but loved the shoe and the feel of it	Neutral	Đánh giá này của khách hàng khó phân biệt cảm xúc khi khách hàng vừa khen ngợi vừa phê bình sản phẩm
Two left found in the box	Neutral	Đánh giá chỉ cung cấp thông tin, không thể hiện cảm xúc

2.3 Thống kê ngữ liệu

Ở phần này, nhóm thực hiện các thao tác tính toán cơ bản trên Python để thống kê số lượng và phân bố của các từ trên tập dữ liệu huấn luyện để có cái nhìn tổng quan và hiểu rõ hơn về ngữ liệu đang sử dụng

Bảng 2.2 Thống kê tổng quan về bộ ngữ liệu

Nội dung	Positive	Negative	Neutral
Số lượng mẫu	35977	35984	18039
Số từ trung bình mỗi câu	28.22	37.09	36.06
Số từ vựng	22068	26560	16608
Số từ trung bình mỗi câu trên toàn bộ tập dữ liệu	33.33		
Tổng số từ trên toàn bộ tập dữ liệu	41982		



Hình 2.2 Phân bố nhãn trên từng lớp của bộ dữ liệu

Dựa trên bảng thống kê và sơ đồ phân phối của các nhãn trong dữ liệu, nhóm rút ra một vài nhận xét cho ngũ liệu như sau:

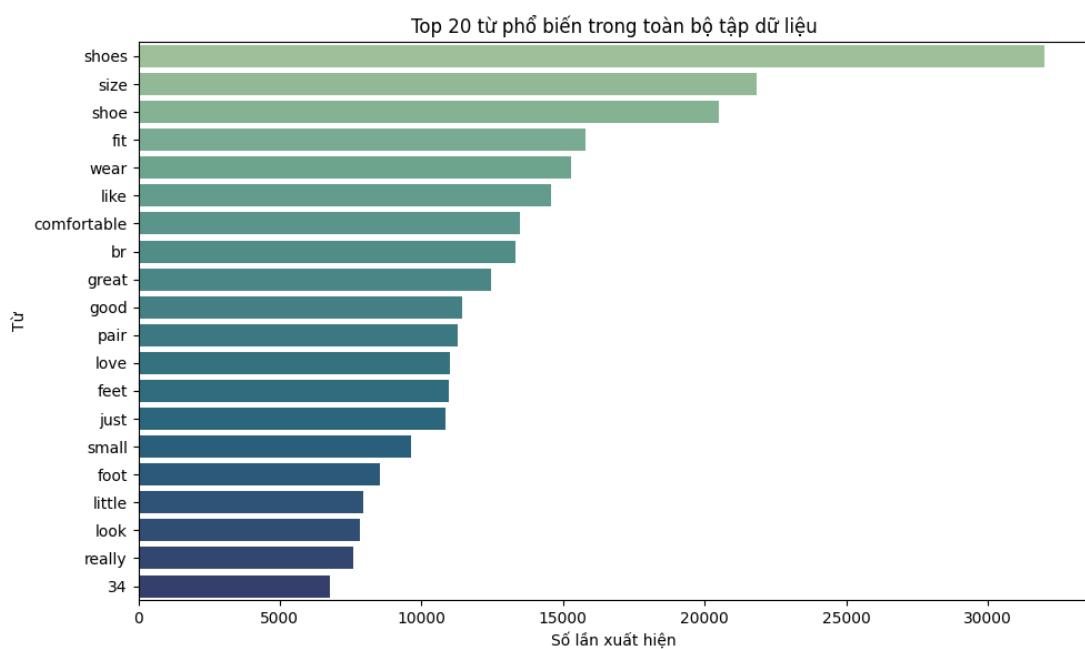
- Số từ trung bình trong lớp Positive thấp hơn so với hai lớp Neutral, Negative, điều này có thể được giải thích rằng đối với một người mua hàng có thái độ hài lòng với sản phẩm, họ sẽ đa số khen và viết một cách ngắn gọn, không trình bày quá nhiều, “không có gì để chê”. Ngược lại với hai nhãn còn lại, khách hàng có xu hướng trình bày về những trải nghiệm chưa hài lòng, không đáng có hoặc trung lập mà mình gặp phải, nhất là lớp Negative, đây vừa là điều có lợi và có hại cho doanh nghiệp, họ sẽ nhận được những sai sót, những gì chưa tốt ở sản phẩm để có thể sửa đổi hoặc hiệu chỉnh, tuy nhiên các đánh giá tiêu cực có khả năng sẽ hướng người dùng đến cửa hàng khác để mua. Đây cũng là thông số quan trọng, vì độ dài văn bản có thể ảnh hưởng phần nào đến cách mà mô hình học và phân loại cảm xúc, nhóm sẽ cân nhắc đến việc chuẩn hóa độ dài các mẫu

- Số từ vựng tổng của tập dữ liệu huấn luyện là 41982, một con số khác lớn về số từ vựng khác nhau mà khách hàng đánh giá sản phẩm, cho thấy được độ đa dạng

của ngũ liệu. Điều này đặt ra một yêu cầu trong sử dụng các kĩ thuật trích xuất đặc trưng phù hợp nhằm giảm chiều dữ liệu và giữ lại thông tin mang tính phân biệt, có ý nghĩa cao

- Về phân bố tỉ lệ, so với các bộ dữ liệu thực tế vốn thường mất cân bằng mạnh thì bộ dữ liệu này tạm chấp nhận về sự cân bằng với tỉ lệ 4 – 4 – 2, tuy nhiên điều này vẫn tiềm ẩn nguy cơ gây ra thiên vị (bias) cho mô hình khi học trên bộ dữ liệu. Vì vậy cần áp dụng các kĩ thuật cân bằng dữ liệu

Đối với các đánh giá, thông thường sẽ có những từ thường xuyên xuất hiện, tùy thuộc vào ngũ cảnh, môi trường hiện tại, ở đây là các đánh giá liên quan đến sản phẩm về giày dép cho nên các từ phổ biến nhất thường sẽ liên quan đến giày dép và đặc tính của nó. Vì vậy để trực quan hóa các từ phổ biến, nhóm sử dụng CountVectorizer để thống kê và in ra 20 từ phổ biến nhất trên toàn bộ tập dữ liệu



Hình 2.3 Top 20 từ xuất hiện phổ biến trong tập dữ liệu

Qua sơ đồ về top 20 từ phổ biến, nhóm rút ra một vài nhận xét như sau:

- Như đã đề cập, các từ ngũ sẽ xoay quanh các chủ đề liên quan đến giày dép, các từ phổ biến xoay quanh chủ đề này
- Ngoài các từ mặc định sẽ xuất hiện nhiều là “shoes”, “shoe” thì các từ như “fit”, “size”, “comfortable” xuất hiện với tần suất cao, cho thấy khách hàng rất quan

tâm đến cảm giác mang và độ phù hợp về kích cỡ của sản phẩm, điều này hoàn toàn hợp lý bởi nếu như cảm giác mang và kích cỡ phù hợp sẽ khiến cho họ ấn tượng và quay lại cửa hàng

- Vì đang xét trên từng từ nên số lượng các từ tích cực chiếm ưu thế trên biểu đồ này, tuy nhiên cần phải xem xét kỹ bởi chúng có thể đi kèm theo các từ phủ định làm cho ngữ nghĩa không giống như suy nghĩ chủ quan

- Ngoài ra có các từ không mang ý nghĩa như “br”, theo nhóm quan sát từ dữ liệu thì đây là kết quả xử lý HTML chưa triệt để khi chưa xóa hết các thẻ
, và 34 có thể đại diện cho size giày, nhóm sẽ thực hiện loại bỏ các từ không mang ý nghĩa và các con số không cần thiết trong quá trình tiền xử lý dữ liệu

2.4 Kiểm tra và giải thích các kết quả gán nhãn

Ở phần này, nhóm sẽ lựa chọn ra 72 mẫu ngữ liệu trong dữ liệu huấn luyện, chia đều ra 3 lớp để thực hiện giải thích kết quả gán nhãn và kiểm tra liệu các kết quả gán nhãn đã phù hợp chưa

Bảng 2.3 Kiểm tra và giải thích gán nhãn của 72 mẫu

STT	Nhãn	Nội dung đánh giá	Giải thích kết quả
1	Positive	Good shoe for office work. They will scuff very easy so be aware	Đánh giá này cho thấy khách hàng cảm thấy sản phẩm vô cùng phù hợp cho công việc văn phòng
2	Positive	Excellent shoes , very confortable and litghweight !	Đánh giá này thể hiện rõ sự yêu thích của khách hàng với các từ ngữ như “comfortable”, “excellent”
3	Positive	Everything about the boot is great	Đánh giá này là sự khen ngợi của khách hàng về mọi mặt của sản phẩm
4	Positive	Slippers fit great. Nice and cozy.	Đánh giá này thể hiện cảm nhận tích cực của khách hàng khi sử dụng sản phẩm với các từ ngữ như “fit great”, “nice”, “cozy”
5	Positive	Fits well.	Đánh giá này ngắn gọn nhưng thể hiện cảm giác mang vừa vặn của khách hàng

6	Positive	True size 9, very comfortable and pretty. These were purchase to wear on campus, during heavy course days, that requires plenty of walking	Đánh giá này thể hiện sự yêu thích của khách hàng với “comfortable”, “pretty” và khách hàng còn mô tả thêm về công dụng cũng như lợi ích thực tế của sản phẩm
7	Positive	I only have worn them once or twice but very comfortable and supported my feet for good length of time. I wish you had WIDE shoes thank you	Đánh giá này thể hiện sự yêu thích của khách hàng và họ khuyến khích tạo ra thêm nhiều sản phẩm tương tự
8	Positive	Good sunglass	Đánh giá ngắn gọn, khen về sản phẩm
9	Positive	Comfortable. I wear slippers all the time at home. These are a bit warm for the summer, but very good overall	Đánh giá này vừa khen ngợi vừa kể về trải nghiệm thực tế khi sử dụng sản phẩm
10	Positive	I have converted my whole family to Tubes and a handful of friends. I love these shoes. The leather ones like these are a little warm but just as comfy as the rest of my tubes. I'm on my 3rd pair in 4 years and	Đánh giá thể hiện rõ sự yêu thích của khách hàng, vị khách này giới thiệu sản phẩm đến gia đình và bạn bè cũng như kể về việc thường xuyên mua lại sản phẩm này cũng như lưu ý về khả năng bị trượt cho khách mua sau

		where them just about every day. The only downside to these shoes is they can be slippery on some surfaces, especially when wet. Be careful!	
11	Positive	I wear a 10.5 ordered 10..fits great! Real fancy! Love the color.. Just gotta wear them in.. Neck is stiff.. First pair of cowgirl boots I have owned!	Đánh giá thể hiện rõ sự yêu thích với “fit great”, “fancy”, “love”
12	Positive	Basic shoes.	Đánh giá ngắn gọn, nói về sự đơn giản của đôi giày, khả năng đánh giá này ở thang đo 4 sao
13	Positive	I bought these for my husband for his birthday. He had been eyeballing them. Yes they are on the more expensive side for being sandals, but they have been well worth the money. He has worn them all spring and summer long in all kinds of weather and conditions. They are	Đánh giá ban đầu có nói về việc giá cao so với mặt bằng chung, tuy nhiên sau đó là sự khen ngợi và trải nghiệm tốt của khách hàng thì chất lượng tương ứng với giá tiền

		really holding up their shape and quality. He pretty much wears them everyday after coming home from work. Wears them all weekend long. He loves them and they are well worth the price.	
14	Positive	I just bought them, at first feels like kind of stretch on the sides, but will see how feels with some use...	Đánh giá cho thấy trải nghiệm ban đầu chưa thực sự quá tốt nhưng họ muốn xem cảm giác thế nào sau một thời gian sử dụng, khả năng thuộc lớp 4 sao
15	Positive	Wore for my wedding. Loved them!	Đánh giá thể hiện rõ sự yêu thích của khách hàng
16	Positive	It's very nice shoes! It's comfortable.	Đánh giá thể hiện rõ sự yêu thích của khách hàng
17	Positive	Love this bag! Perfect size for my wallet, cellphone, keys and sunglasses. The lining is striped so I can find things easier than a solid dark color.	Đánh giá thể hiện rõ sự yêu thích của khách hàng về công năng sản phẩm
18	Positive	Nice looking shoe. Comfortable	Đánh giá thể hiện rõ sự yêu thích
19	Positive	The hard case is large enough to protect the glasses and provides	Đánh giá vừa khen ở một khía cạnh, vừa chê ở một khía cạnh, tuy nhiên họ sử dụng “the only negative...” tức là trong các

		adequate outside carry potential. The only negative comment I have is the zipper does not open or close easily because the zipper material is too flexible to allow closure until both zipper sides are manually (by hand) melded together to allow the zipper to begin to open or close.	điểm cộng chỉ có duy nhất một điểm trừ, vậy nhãm tích cực cho trường hợp này là hợp lý, khách hàng muốn góp ý về sản phẩm cho doanh nghiệp
20	Positive	Perfect carry on or weekend bag	Đánh giá thể hiện rõ sự khen ngợi
21	Positive	Excellent boot! Hardly no break in period (would definitely recommend and would buy again).	Đánh giá thể hiện rõ sự yêu thích và khen ngợi sản phẩm, ngoài ra còn đề xuất mọi người mua và họ sẽ mua lại
22	Positive	I can run in these if I wanted to, and keep my feet warm during winter	Đánh giá thể hiện sự khen ngợi đối với sản phẩm
23	Positive	Love the sunglasses. Love the look. Love the Polarized lenses. Love the price. However, the lenses have popped out	Đánh giá cho thấy khách hàng rất yêu thích các sản phẩm, tuy nhiên khi gấp ván đề về lỗi sản phẩm, họ đã có găng liên lạc với bộ phận hỗ trợ nhưng không nhận được phản hồi và đánh giá của họ từ 5 sao còn 4 sao nhưng vẫn trong thái độ tích cực

		<p>twice. Once from a 5 foot fall from my head on a carpeted floor. Another sitting in the case in my gym bag. I emailed customer support and never received a response. Would have been a 5 star but loses a star for poor customer service.</p>	
24	Positive	<p>Good sandals for around the pool but not a pair I'd wear the whole day through. The footbed, as the listing called it, is very cushy and the rubber feels like memory foam. Recommended.</p>	<p>Đánh giá cho thấy sản phẩm phù hợp và nhận được sự khen ngợi của khách hàng, tuy họ có đề cập rằng sản phẩm không phải loại mà họ có thể mang cả ngày, tuy nhiên họ vẫn đề xuất cho các khách hàng khác nên mua</p>
25	Negative	<p>I have had the Patricia II wedge in black for about 1 year & wore them regularly in season. When I saw the Patricia at a good price in navy (from 6pm), I purchased them because I thought they would fit just like my Patricia</p>	<p>Đánh giá khá dài và chi tiết về trải nghiệm của khách hàng, họ đã mua mẫu dép khác (Crocs) vì nghĩ mẫu này giống mẫu mà họ đang sử dụng, tuy nhiên gặp vấn đề về kích thước mẫu mới, size 9 thì quá rộng với họ, size 8 thì vừa chân trái nhưng quá chật chân phải (điều này khá dễ hiểu vì phần lớn đi thử giày sẽ thử ở chân bên phải trước, vì nếu chân bên phải vừa chắc chắn chân bên trái cũng sẽ vừa), từ đó khách hàng nghĩ rằng việc chênh lệch về</p>

		<p>IIs. I was wrong, and paid the price with return shipping that 6pm doesn't pay. Crocs website says that crocs aren't suppose to fit like other sandals - they are suppose to be looser & thus more comfortable - I normally wear an 8-1/2, so have now tried both an 8 & a 9 in the Patricia shoe (I have an 8 in the Patricia II). The Patricia 9 swims on my feet & they would be a hazard to walk around in. The size 8 fits my left foot (which is my wider foot) but is too narrow on my right foot. When I placed the shoes sole to sole, I did notice a slight difference in the width, which, apparently, my foot notices too. I can only conclude a</p> <p>độ rộng giữa hai kích thước là lỗi sản xuất và khách hàng cũng đang phân vân không biết rằng liệu do đôi dép này hay do khuôn sản xuất của toàn bộ mẫu dép. Trong đánh giá này không thể hiện quá rõ sắc thái của khách hàng, tuy nhiên qua sự chia sẻ về trải nghiệm của mình cũng như sự hoang mang của khách hàng thì khả năng đây là đánh giá 2 sao, cũng được xem là hợp lý khi gán nhãn Negative</p>
--	--	---

		<p>manufacturing defect.</p> <p>But, it is this shoe specifically or the form for this shoe? (others have written the shoe is narrow).</p> <p>Consequently, if you have a wider foot, order the Patricia II instead of this one & if you are a 1/2 size, order down, not up</p>	
26	Negative	<p>Width not right and size too small if width had been just little wider and ordered size larger would have been good. Loved the shoe look</p>	<p>Đánh giá phần lớn phê bình về chiều rộng và kích thước sản phẩm quá nhỏ, sẽ tốt hơn nếu chiều rộng được tăng thêm và đặt size lớn hơn, tuy nhiên họ vẫn có một điểm cộng yêu thích ngoại hình sản phẩm</p>
27	Negative	<p>I received these shoes and they weren't the same as the picture described them, they were a different color. When i tried to return them, the shipping wasn't paid for. So i had to pay \$20 for shipping. A waste of time and money. I dont recommend</p>	<p>Đánh giá cho thấy trải nghiệm vô cùng tệ của khách hàng, sản phẩm họ nhận về so với hình ảnh tham khảo khác biệt về màu sắc, khi có nhu cầu trả hàng thì bên giao hàng không chấp nhận phí vận chuyển mà họ phải tự bỏ tiền túi để trả, họ cho rằng phí phạm thời gian và tiền bạc, khuyên mọi người không nên mua</p>

		anyone to buy from TheSmartBuy.	
28	Negative	Usually love Ethnies product. In this case the raised arch area of one shoe is too far back and raised to the point of discomfort. The other shoe fits fine.?? Had to wear em before I figured it out so im stuck with em. Also, they do run a bit narrow(or at least one shoe did) ;)	Đánh giá cho thấy sự thất vọng của khách hàng đối với sản phẩm của nhãn hàng yêu thích, thiết kế của sản phẩm làm cho họ cảm thấy khó chịu và họ phải chấp nhận giữ lại sản phẩm vì họ đã mang thử trước khi nhận ra vấn đề
29	Negative	Could not get my foot into the shoe. Was disappointed and returned them	Đánh giá thể hiện rõ sự khó chịu và thất vọng của khách hàng, họ không thể đưa chân mình vào bên trong giày
30	Negative	Pros: They are very comfortable. I wear them all day long, every day with no complaints. They are also well ventilated to keep your feet cool. This makes them a little more prone to your socks getting wet from a puddle but not	Đánh giá này khá chi tiết khi khách hàng liệt kê rõ ưu và nhược điểm của sản phẩm, tuy nhiên phần nhược điểm khá nhiều vấn đề xảy ra, và khách hàng phải dùng đến cụm "...this is the worst quality pair I've ever owned", cho thấy sự thất vọng và khó chịu của khách hàng về chất lượng sản phẩm, là đôi giày có chất lượng tệ nhất từ trước đến nay mà họ gặp

		a big deal. Cons: They aren't built to last. 4 months after I bought them they are junk. The rubber bottoms on both heels fell off. The side stripes are falling off, and the toe soles are coming off now too. I've been wearing Asics for the last 10 years and this is the worst quality pair I've ever owned.	
31	Negative	These look great, the lenses are nice color and mirrored, and the matte finish is really nice. The good stops there. As soon as I put these on in my car I immediately notice big reflections that were very bright and hard to ignore. I thought that maybe I had left a sticker or some kind of plastic cover on the inside because it was crazy. The INSIDE of these	Đánh giá này chỉ khen ngợi ở câu đầu tiên, còn lại phía sau đều là sự chê bai, phê bình về chất lượng và trải nghiệm không có đáng có khi dùng sản phẩm, cuối cùng họ phải trả lại và dùng từ “useless” cũng như khuyên mọi người không nên mua, thể hiện rõ thái độ tiêu cực

		are so reflective that it takes the light from the SIDE and BEHIND you and point it right back at your eyes. I seriously couldn't believe that these glasses meant to block out sunlight were actually reflecting it back into my eyes. I've bought a ton of stuff on Amazon and these might be the first I return. Truly truly useless.. DON'T BUY	
32	Negative	I was expecting it to be great, but they do run small. It cuts into my ankle and I can't wear them to work as it was its purpose.... Highly suggest buying a size up.	Đánh giá thể hiện khách hàng mong đợi chất lượng sẽ tốt hơn, tuy nhiên sản phẩm quá nhỏ, gây chấn thương trong quá trình sử dụng cho khách hàng, tuy nhiên họ không tẩy chay mà gợi ý mua size lớn hơn
33	Negative	The quality is awful, shoes started falling apart the first week I had them. Not the first fakes (or at least very poor quality items)	Đánh giá thể hiện chất lượng sản phẩm quá tệ khiến khách hàng cho rằng đây là một trong những sản phẩm tệ nhất từng nhận được từ Amazon

		that I have received from Amazon	
34	Negative	Not very flexible. Top strap is stiff, not pliable at all. After a week of use, there was little sign that it will relax any. Plastic from strap has sharp edge that irritated my foot top	Đánh giá về chất lượng sản phẩm không tốt, thường xuyên gặp các vấn đề với sản phẩm
35	Negative	My mom always had these shoes for my sister and I when we were babies (or a very similar style), so I thought they'd be great everyday shoes for my one year old. She already had shoes she was about to grow out of, so I got the next size up in advance so we'd have them when she was ready to wear them. I tried them on her when they arrived to check on the width issue based on other reviews, and it seemed she'd be able to wear	Đánh giá chi tiết về trải nghiệm mua hàng không đáng có của khách hàng, họ đã mua trước sản phẩm có kích thước lớn hơn dự tính cho trẻ em trong nhà khi lớn lên có thể mang vừa, ở lần đầu thử thì sản phẩm vừa với chân của em bé, sau 6 tuần thì kích thước không còn vừa và sản phẩm bị bung đường may, họ muốn thực hiện trả hàng như chính sách đổi trả cho sản phẩm chỉ là 30 ngày và họ mặc định nghĩ rằng các sản phẩm sẽ có thời gian đổi trả là 90 ngày, họ cảm thấy lãng phí khi mua đôi giày chưa từng sử dụng và bị hỏng

		<p>them. In the six weeks since then, her feet have not just grown in length but width as well, because now that her feet are long enough for the shoes, there is absolutely no way I can velcro these shut without squeezing her poor little feet! In addition to that, the first time I even tried to get them on her, the loop that holds the strap in the middle popped right open (the seam came apart). I immediately got on my computer to print a return receipt only to discover the return policy is only 30 days on these (I'll admit I'm used to returning items to stores that have 90 day policies). Now I have a pair of damaged, too-small shoes that we will ever even use. What a waste!!</p>	
--	--	---	--

36	Negative	is big for me ...definitely return it back.shipping was good and fast	Đánh giá cho thấy sản phẩm quá lớn so với khách hàng và họ đã trả hàng, tuy nhiên họ có lời khen cho bên vận chuyển. Đây vẫn được xem là đánh giá mang sắc thái tiêu cực vì sản phẩm họ nhận được không phù hợp
37	Negative	I just got the shoe's and wore them around the house for over an hour. I will use them as junker shoe's because that's really all they are good for. I have had several pairs of New Balance running shoes (Other pair is 750V1 - I really like those), these are garbage. They fit is tight and small, and they are narrow. No real arch support. It is almost like they took a tennis shoe and covered it with a running show shell? The quality and construction is pretty poor and doubt they will last very long. Basically the shoes are	Đánh giá thể hiện rõ thái độ chê bai của khách hàng với cụm “these are garbage”, họ cũng nêu chi tiết về những điểm của sản phẩm gây ra cảm giác khó chịu và không thoải mái

		uncomfortable, don't waste your money. I don't want to go through the hassle of returning them and I will wear them (but working around the house outside). The one saving grace is aesthetically they look good	
38	Negative	These shoes have a good look but don't last long. I got mine Mar 11, so I'm at about month five with these and casual wear. Both shoes are coming apart along the outer edge where the canvas meets the rubber. I was a disappointed but now the left one started to make noise when I walk, hence the one star. My right shoe also got a puncture hole in it from something, and I have no idea how just glad it didn't get	

		through to my foot. I have multiple other pairs of Merrells that have been durable but these seem to not be made as well	
39	Negative	Bad lens	Đánh giá ngắn gọn, thể hiện sự chê bai
40	Negative	bad quality the color went off after 2 weeks	Đánh giá ngắn gọn, thể hiện sự chê bai chất lượng, phai màu trong 2 tuần
41	Negative	Bought two pairs of these shoes for my twins. They were trashed in a day. Fit was stiff & small. Don't waste your time or money	Đánh giá khá ngắn gọn và thể hiện rõ thái độ chê bai về độ hỏng của giày, mang cứng và chát, khuyên mọi người không nên tốn thời gian, tiền bạc
42	Negative	These were huge! I have other nw shoes and these were so much bigger. Have to return	Đánh giá cho thấy sản phẩm quá to, khách hàng phải trả lại hàng
43	Negative	These are SO cute but they did not fit my daughter the way I had hoped. The strap does not come off so they are a little difficult to get on. My daughter has wide feet and so they did not fit very well. Maybe would	Đánh giá mang sắc thái thất vọng và tiếc nuối của khách hàng, mặc dù vẻ ngoài của sản phẩm dễ thương nhưng không vừa với con gái của họ như mong đợi, gặp khó khăn khi xỏ chân vào và bé gái gặp trải nghiệm không đáng có khi chân phòng rộp lúc mang vào ngày đầu tiên

		have fit better if I got one size bigger than she wears. She got blisters the first day she wore them. :(
44	Negative	It's very cute but there was a seam sewn right on the front and the block spots began to rub white. I returned the item	Đánh giá của khách hàng về chi tiết sản phẩm không tốt và đã trả hàng
45	Negative	These shoes are great except for the soles - they're made up of several small patches of rubber glued to a soft foam sole. On each shoe, two of the rubber patches are peeling off, I've had to re-glue them a handful of times. Also, the foam part gets pretty torn up on rough terrain. After a couple of hikes in these I've decided to use them just as regular sneakers for every day, casual use. I love them for casual use,	Đánh giá khá dài và chi tiết, khách hàng nêu ra trải nghiệm của mình về sản phẩm ở hai giai đoạn là lúc mới mua và sau 6 tháng sử dụng. Ở giai đoạn mới mua, khách hàng đã mô tả một vài chi tiết về khả năng hỏng hóc của đôi giày, tuy nhiên họ vẫn có trải nghiệm tốt khi mang trong lúc đi ra ngoài. Tuy nhiên sau khi mang 6 tháng, các chi tiết hỏng dần xuất hiện nhiều và rõ hơn, họ có nói rằng mang vẫn thoải mái nhưng chất lượng đi xuống khá nhanh và thay đổi đánh giá từ 3 sao xuống còn 1 sao. Qua đánh giá thì khách hàng vẫn cho rằng việc mang đôi giày này khá thoải mái, tuy nhiên tốc độ xuống cấp về chất lượng là quá nhanh, họ cũng giảm mức đánh giá xuống, cho thấy thái độ khá tiêu cực

		<p>they're comfortable and the uppers seem pretty durable. I've worn them pretty much every day for the past 2.5 months and they've held up fine. 6 Month update - Patches of rubber have torn off, others refuse to stay glued, and there is a hole in the upper portion on each shoe. Additionally, there's a hole through the sole at the heel of my left shoe, AFTER 6 MONTHS! Soles should not wear out in six months! They're still comfortable but they've worn out way too fast. I'm decreasing my rating from 3 stars to 1 star. I'm thankful I didn't pay very much for these</p>	
46	Negative	Received these shoes and had to turn around and send them right	Đánh giá không tốt từ khách hàng về phần đế giày, họ cho rằng là dấu hiệu không tốt về chất lượng và ngay lập tức trả hàng

		back. The footbed liner was coming up on both shoes before I even wore them; not a good sign of quality, or craftsmanship	
47	Negative	Were comfortable the 1st time, but seem to get tighter everytime I wore them 😞 recd many compliments on them but can no longer wear them	Đánh giá mang tính thất vọng của khách hàng, lần đầu mang thì họ cảm thấy thoải mái nhưng dần dần càng bị chật, và mặc dù sản phẩm này được nhiều lời khen nhưng họ không thể mang được nữa
48	Negative	Hurts my feet. Like wearing razor blades. The material is too scratchy and I can't even wear them. Even tried with hose on but they are aweful. I wore them to work one day, I stay at a desk mostly and couldn't even tolerate wearing them sitting down. They are attractive but OUCH!	Đánh giá vô cùng tệ về trải nghiệm khi sử dụng sản phẩm, khách hàng bị đau, kể cả có sử dụng vớ và họ cũng không di chuyển nhiều nhưng vẫn không thể chịu nổi, ngay cả khi ngồi xuống.
49	Neutral	They began to split alone the mesh material after a month	Đánh giá cho thấy chất lượng sản phẩm đi xuống nhưng khách hàng vẫn yêu thích đôi giày và cảm giác khi mang, mang hai sắc thái cảm xúc

		but loved the shoe and the feel of it	
50	Neutral	Nice looking shoe, okay for short-term wear. Much narrower than other size 11's that I have - tight fit!	Đánh giá cho thấy giày đẹp, phù hợp để mang nhưng hẹp và chật, mang hai sắc thái cảm xúc
51	Neutral	The shoe was the EXACT color I was looking for but... You get exactly what you pay for. Poorly made shoe. Fit comfy but the heel feels like it wobbles a little when I walk. Was going to wear them to an outside wedding, decided against it. This shoe is probably more for a day at work.. or church. No where there would be a lots of walking around	Đánh giá cho thấy ban đầu khách hàng yêu thích bởi màu sắc của giày, tuy nhiên sau đó là phản chê bai về chất lượng, những gì mà họ nhận được đúng với số tiền bỏ ra, tức ám chỉ giá tiền của sản phẩm thì rẻ, đi đôi với đó là chất lượng không tốt, các câu sau đó nói về trải nghiệm không mấy thoải mái khi sử dụng và họ thấy chỉ phù hợp khi đi làm, không phải đi lại nhiều. Đây là đánh giá mang đúng sắc thái trung tính, bỏ số tiền bao nhiêu thì nhận lại chất lượng bấy nhiêu
52	Neutral	I loved how comfortable these shoes felt when I put them on. However, after a while, my big toe rubs against the toe cutout. To prevent	Đánh giá vừa thể hiện sự yêu thích về độ thoải mái của sản phẩm, vừa nói về điểm trừ của sản phẩm, và khách hàng cũng tự tìm cách khắc phục vấn đề về điểm trừ đó, cuối cùng họ cũng nói về cửa hàng yêu thích của họ nhưng có một chút thất vọng

		the rubbing, I wear the shoe with a Band Aid across my toe and it helps. Donald Pliner is my go to brand for shoes. I am a little disappointed with this pair	với đôi giày hiện tại, đây là sắc thái trung tính
53	Neutral	These shoes are cute as several people have said, however, they are cut low and do not lace all the way up to your ankle. I was concerned that they would slip when walking the golf course, so I returned them. Too bad.	Đánh giá của khách hàng cho rằng kiểu dáng đôi giày dễ thương như bao người khác đánh giá, tuy nhiên thiết kế của giày khả năng cao sẽ gây tuột khi đi bộ trên sân golf nên họ đã trả lại, “too bad” trong trường hợp này mang nghĩa là thật tiếc, tức là họ vẫn thích nhưng vì lý do thiết kế nên trả lại, cũng được xem là sắc thái trung tính
54	Neutral	Fitted smaller than expected	Đánh giá cung cấp thông tin, khách hàng đánh giá rằng nhỏ hơn so với mong đợi, không thể hiện rõ yêu thích hay chê bai
55	Neutral	The shoe looks amazing but the width is too small .i really like the look	Đánh giá của khách hàng vừa khen ngợi về ngoại hình sản phẩm, vừa có sự chê bai về độ rộng
56	Neutral	for the price, I am disappointed that the paint or whatever the outside layer is, is peeling. They also	Đánh giá này thể hiện cả hai sắc thái tích cực và tiêu cực, khách hàng thất vọng về giá tiền phải bỏ ra mà chất lượng không tương xứng, tuy nhiên họ khá thích về màu sắc. Đây là đánh giá đầu tiên mà

		scratched SO easily. But the pink lenses have been awesome	nhóm cho rằng nên thay đổi nhãn từ Neutral thành Negative
57	Neutral	Not quite what I expected. Sized small. Still like them though.	Đánh giá này thể hiện cả hai sắc thái, khách hàng cho rằng không như họ mong đợi vì size nhỏ tuy nhiên họ vẫn thích chúng
58	Neutral	nice looking...hopefully will wear well	Đánh giá này thể hiện rõ sự yêu thích của khách hàng về ngoại hình và họ mong đợi sẽ mang tốt, đánh giá này mang tính xây dựng và có thiên hướng nghiêng về Positive, nên thay đổi nhãn từ Neutral thành Positive
59	Neutral	The size is too strange	Đánh giá ngắn gọn, khách hàng cảm thấy kích thước sản phẩm khá lạ, không thể hiện rõ cảm xúc
60	Neutral	ok for the present time	Đánh giá ngắn gọn, ổn trong thời điểm hiện tại, cũng không thể hiện quá rõ cảm xúc
61	Neutral	A bit uncomfortable, whether due to my ankle bones or the moccasin style, or just needing to be broken in more; these develop pressure points on the side of my ankles where the throat of the shoe rubs against them. The footbed is great	Đánh giá vừa có khen vừa có chê, cảm giác mang không thoải mái, khách hàng cho rằng có thể do cấu tạo mắt cá chân của họ hoặc cần mang nhiều lần để thoải mái hơn, nhưng vẫn khen phần đế trong

62	Neutral	<p>These are very attractive shoes. I wear an 8 or 8.5 normally. The 8 was too small. I returned them due to the fit and the thin soles. However they are so cute I'm thinking about trying the 8.5. Not much to lose at this price point...</p>	Đánh giá này phần lớn là sự yêu thích của khách hàng về đôi giày, vấn đề duy nhất là họ cảm thấy không vừa và đế giày khá mỏng nhưng họ vẫn đang cân nhắc mua size lớn hơn, nên thay đổi từ Neutral sang Positive
63	Neutral	<p>coming apart where the toe meets the body (by the ball of the foot). I expected them to hold up better than they have. Other than that no complaints. They are comfortable for a hiking shoe</p>	Đánh giá vừa có khen vừa có chê, điểm trừ là chất lượng chi tiết, không bền tuy nhiên họ vẫn cảm thấy thoải mái khi đi bộ đường dài, mang hai sắc thái
64	Neutral	Cute but smaller than expected for the \$\$	Đánh giá cho thấy khách hàng mong đợi size sẽ lớn hơn và khen về sự dễ thương của đôi giày
65	Neutral	A little smaller than my other sandals	Đánh giá mang tính cung cấp thông tin
66	Neutral	This shoe is not as comfortable as I expected it to be.	Đánh giá mang tính trung lập, họ cảm thấy không thoải mái như mong đợi nhưng cũng không bị khó chịu
67	Neutral	Two left found in the box ..	Đánh giá mang tính cung cấp thông tin

68	Neutral	needs half sizes I usually wear an 8 the 8 was way to small as in my toes were hanging over the edge. the 9 is way too big.	Đánh giá mang tính góp ý và đề xuất của khách hàng cho doanh nghiệp nên làm thêm size nữa, size 8 thì chật và size 9 thì quá rộng
69	Neutral	used to really love these. They're still comfy and great, I could hike 3-5 miles in these. However I went on my longest hike of 10 miles and my toes were ACHING. I ordered half a size bigger and Never had a problem with them until my longest hike. Really wish I could've worn these forever but for long hikes, not for me!	Đánh giá vừa có khen vừa có chê, khách hàng khá thích đôi giày này, mang lại cảm giác dễ chịu, tuy nhiên khi thực hiện đi bộ đường dài thì họ gặp tình trạng đau nhức chân và cho rằng không phù hợp, họ mong rằng có thể mang đôi này ở bất cứ trường hợp nào nhưng thực tế không phù hợp với leo núi và đi đường dài, vẫn mang sắc thái trung tính
70	Neutral	good looking shoe, No support	Đánh giá ngắn gọn, có khen về ngoại hình nhưng không ủng hộ, không thể phân biệt rõ cảm xúc
71	Neutral	his is my second pair of Moabs that I ordered after my 3+ year old pair are starting to wear out. I loved my original pair.	Đánh giá khá dài và chi tiết, đây là đôi thứ 2 mà khách hàng mua, khách hàng rất thích đôi mua đầu tiên vì các trải nghiệm và công năng của sản phẩm. Tuy nhiên đến đôi thứ 2 thì họ cảm thấy không ổn, cảm giác mang vẫn thoải mái nhưng cầu

		<p>They were extremely comfortable, water proof, breathable enough (but the waterproofing does trap some heat), grippy, and rugged. That said, something about this new pair seems off. They are mostly comfortable, but the heel support doesn't quite match between the two shoes. One seems slightly higher and more stiff than the other. There is also some odd clicking near my right front toe. I think it has to do with the waterproofing on the inside of the shoe. My original pair was my favorite shoe of all time. This new pair isn't living up to that standard. I definitely was expecting more quality control at this price. As a long-time</p> <p>tạo của đôi giày bên thấp bên cao, không đúng chất lượng, với mức giá đó thì họ mong rằng chất lượng sẽ tốt hơn và cảm thấy có hơi thất vọng, và họ sẽ cập nhật thêm sau thời gian sử dụng. Đây là đánh giá mang xu hướng góp ý, mang cả hai sắc thái cảm xúc, mong muốn có sự kiểm định chất lượng của cửa hàng, doanh nghiệp</p>
--	--	--

		previous owner of this shoe, I was let down with the new purchase. I plan on updating to see if things even out or get worse as I continue to break them in.	
72	Neutral	I wear a size 7 in all my shoes but this one was too large. Hurt the back of my feet because they were too big.	Đánh giá mang tính chê bai của khách hàng về size của giày, mang đến trải nghiệm không tốt, nên thay đổi sang sắc thái tiêu cực

Sau khi nhóm thực hiện quá trình trích xuất ra các ngữ liệu và thực hiện giải thích cũng như kiểm tra về kết quả gán nhãn cho từng ngữ liệu thì nhóm rút ra được một vài nhận định như sau:

- Các mẫu trong hai lớp Positive và Negative được gán nhãn hoàn toàn trùng khớp và phù hợp, bởi lẽ các sắc thái được thể hiện qua các câu từ trong đánh giá nằm ở mức khá rõ ràng đến rất rõ ràng của người mua ở 1 - 2 sao và 4 - 5 sao
- Các đánh giá mà nhóm cho rằng có sự sai sót và cần chuyển đổi nhãn sẽ được tô đỏ toàn bộ dòng, trong đó toàn bộ là những đánh giá ở lớp Neutral, có những đánh giá nên chuyển sang Positive và có những đánh giá nên chuyển sang Negative, đây là điều dễ hiểu bởi trong các đánh giá 3 sao, có những khách hàng gặp những trải nghiệm không đáng có về sản phẩm nhưng họ không muốn làm ảnh hưởng lớn đến đánh giá toàn cục của cửa hàng nên họ vẫn đánh giá ở mức 3 sao, tương tự cũng có những trường hợp các sản phẩm mang lại trải nghiệm tốt và phù hợp nhưng với khách hàng khó tính thì họ vẫn chưa cảm thấy xứng đáng để đánh giá từ 4 đến 5 sao
- Trong đó thì số lượng đánh giá sai sót về nhãn là 4, tỉ lệ trên tổng 72 câu là 5.56% và tỉ lệ trên tổng 24 câu cho lớp Neutral là 16.67%. Như vậy có thể kết

luận rằng vẫn còn một vài mẫu có nhãn không khớp với thực tế nhưng hầu hết vẫn đảm bảo được sắc thái chính mà đánh giá mang lại, và vẫn phù hợp để thực hiện huấn luyện cho các mô hình.

CHƯƠNG 3: PHƯƠNG PHÁP

3.1 Kĩ thuật trích xuất đặc trưng

3.1.1 Term Frequency – Inverse Document Frequency (TF-IDF)

TF-IDF được dùng để đo lường mức độ quan trọng của một từ đối với một tài liệu cụ thể. Nó được tính bằng tích của 2 phần là term frequency (TF) và inverse document frequency (IDF)

TF được định nghĩa là thước đo về mức độ thường xuyên của một từ t xuất hiện trong một tài liệu d cụ thể. Công thức:

$$TF(t, d) = \frac{f_{t,d}}{N_d}$$

Nếu một từ xuất hiện nhiều lần trong một tài liệu, nó có tầm quan trọng cao hơn trong tài liệu đó, tuy nhiên TF sẽ không xét đến việc từ đó có quan trọng trong toàn bộ tập dữ liệu hay không, mà đó là vai trò của IDF

IDF được định nghĩa là thước đo đánh giá mức độ quan trọng của một từ trên toàn bộ tập dữ liệu. Nó được dùng để giảm trọng số của các từ phổ biến như mạo từ vì những từ này xuất hiện nhiều nhưng không mang lại nhiều thông tin cho việc phân biệt và dự đoán. Công thức:

$$IDF(t) = \log\left(\frac{N}{DF(t)}\right)$$

Trong công thức thì t là từ cần tính, N là tổng số tài liệu trong tập dữ liệu và $DF(t)$ là số lượng tài liệu có chứa từ t

Ý tưởng của IDF rất rõ ràng, xuất hiện càng nhiều trong các tài liệu thì ít mang lại thông tin quan trọng, ngược lại, xuất hiện trong ít tài liệu thì khả năng mang lại thông tin quan trọng là cao hơn

Cuối cùng kết hợp cả hai phần lại, ta có được TF-IDF, vừa đánh giá tầm quan trọng cục bộ của một từ trong một tài liệu, vừa điều chỉnh tầm quan trọng toàn cục trong toàn bộ dữ liệu

$$TF-IDF(t, d) = TF(t, d) \times IDF(t)$$

Ví dụ về cách tính của TF – IDF

Bảng 3.1 Ví dụ về cách tính của kĩ thuật TF - IDF

Đánh giá	TF('shoe')	TF-IDF('shoe')
Nice looking shoe. Comfortable	$\frac{1}{4} = 0.25$	$0.25 \times 0.176 = 0.044$
This shoe is not as comfortable as I expected it to be.	$\frac{1}{12} = 0.083$	$0.083 \times 0.176 = 0.0146$
Two left found in the box ..	$\frac{0}{6} = 0$	0
$IDF('shoe') = \log(\frac{3}{2}) = \log(1.5) = 0.176$		

```

good shoe office work stuff very easy so aware
Word  TF-IDF Value
0    aware   0.498937
1    so      0.186992
2    easy    0.361896
3    very    0.176009
4    stuff   0.432408
5    work    0.279014
6    office  0.453347
7    shoe    0.190003
8    good   0.218399

```

Hình 3.1 Ví dụ thực tế về giá trị TF-IDF trên một mẫu trong bộ ngữ liệu

Ưu điểm:

- TF-IDF dễ tính toán, các công thức vô cùng đơn giản và dễ triển khai
- Khác với kĩ thuật trích xuất đặc trưng Bag-of-Words (BoW), TF-IDF có khả năng xác định các từ quan trọng, điều này rất hữu ích để hiểu được nội dung chính của toàn bộ văn bản
- Khả năng phân biệt tốt giữa các từ phổ biến và các từ ít gặp

- Khả năng mở rộng tốt khi TF-IDF có thể xử lý các bộ dữ liệu rất lớn chứa số lượng tài liệu không lồ

Nhược điểm:

- Các từ rất hiếm có thể gây ra vấn đề vì điểm số IDF của chúng có thể bị đánh giá sai lệch cao, khiến chúng có thể mang ý nghĩa quan trọng hơn so với ý nghĩa thực tế mà chúng mang lại
- Cũng như BoW, TF – IDF sẽ gần như không hiểu được ý nghĩa hay ngữ cảnh của từ được sử dụng vì mô hình chỉ đo tần suất xuất hiện của từ mà không quan tâm thứ tự hay các từ đứng trước hay sau nó
- Khó khăn trong diễn giải các từ đồng nghĩa hoặc có nghĩa gần giống nhau, ví dụ như happy và glad đã nêu ở trên, mô hình sẽ gặp khó khăn khi nhận diện những từ này, từ đó có thể dẫn đến các điểm số sai lệch

3.2 Các mô hình sử dụng

3.2.1 Multinomial Naïve Bayes

Multinomial Naïve Bayes (MNB) là mô hình xác suất biến thể của thuật toán Naïve Bayes, được thiết kế để xử lý các đặc trưng rời rạc (discrete feature) như số lần xuất hiện của từ trong văn bản. Ở đồ án này, nhóm sử dụng MNB để ứng dụng vào bài toán phân loại cảm xúc bằng cách kết hợp với kỹ thuật trích xuất đặc trưng TF – IDF

Mô hình dựa trên định lý Bayes với giả định rằng các đặc trưng là độc lập với nhau

$$P(y | x) \propto P(y) \cdot \prod_{i=1}^n P(x_i | y)$$

Trong đó:

$x = (x_1, x_2, \dots, x_n)$ là văn bản đầu vào gồm các từ hoặc đặc trưng

$P(y)$ là xác suất tiên nghiệm của lớp y

$P(x_i | y)$ là xác suất có điều kiện của từ x_i xuất hiện trong lớp y

Và xác suất có điều kiện trong công thức trên được ước lượng với Laplace Smoothing như sau:

$$P(x_i | y) = \frac{N_{i,y} + \alpha}{N_y + \alpha \cdot V}$$

Trong đó:

$N_{i,y}$ là số lần từ x_i xuất hiện trong tất cả các văn bản thuộc về lớp y

N_y là tổng số từ trong lớp y

V là kích thước từ vựng

α là hệ số của Laplace Smoothing và thường mang giá trị 1

Mã giả cho mô hình MNB

Algorithm 1: Multinomial Naive Bayes (Text Classification)

Input: Training data $D = \{(x_1, y_1), \dots, (x_N, y_N)\}$

Output: Predicted class labels

```

1 foreach class  $y \in \mathcal{Y}$  do
2   Compute prior probability  $P(y) = \frac{\# \text{ docs in class } y}{N}$ ;
3   foreach word  $x_i$  in vocabulary do
4     Count  $N_{i,y} \leftarrow$  number of times  $x_i$  appears in class  $y$ ;
5     Compute smoothed likelihood:
6       
$$P(x_i | y) = \frac{N_{i,y} + \alpha}{N_y + \alpha \cdot V}$$

7   foreach new document  $x$  do
8     foreach class  $y$  do
9       Compute score:
10      
$$\log P(y | x) \propto \log P(y) + \sum_i f_i \cdot \log P(x_i | y)$$

11   Assign label  $y^* = \arg \max_y \log P(y | x)$ ;

```

Hình 3.2 Mã giả của mô hình Multinomial Naïve Bayes

Ưu điểm:

- Thuật toán đơn giản, nhanh và dễ dàng triển khai
- Hiệu quả cao với dữ liệu văn bản và đặc trưng rời rạc như TF – IDF
- Dễ hiểu và dễ giải thích ý tưởng cũng như kết quả của mô hình

Nhược điểm:

- Vì thuật toán giả định rằng các đặc trưng là độc lập với nhau, điều này là hoàn toàn phi thực tế đối với các dữ liệu văn bản vì các từ thường có mối liên quan với nhau
- Không xét đến ngữ cảnh, cú pháp hay thứ tự xuất hiện của các từ
- Hiệu suất sẽ giảm nếu số từ vựng quá lớn hoặc dữ liệu không được làm sạch tốt

3.2.2 Logistic Regression

Logistic Regression (LR) là thuật toán học máy có giám sát (supervised machine learning algorithm) được sử dụng cho các bài toán phân loại, thường là bài toán phân loại nhị phân, tuy nhiên có thể mở rộng thành phân loại đa lớp. LR được thiết kế để dự đoán xác suất của một mẫu thuộc về một nhãn cụ thể thông qua hàm sigmoid. Ở đồ án này, nhóm sử dụng LR để ứng dụng vào bài toán phân loại cảm xúc bằng cách kết hợp với kỹ thuật trích xuất đặc trưng là TF – IDF

Về nguyên lý hoạt động, LR học dựa trên hàm tuyến tính:

$$z = w^T x + b$$

Trong đó:

x: Là vector đặc trưng đầu vào, đối với bài toán này là văn bản đầu vào sau khi được biểu diễn bằng TF – IDF

w: Là vector trọng số

b: Là độ lệch (bias)

Sau đó, mô hình thực hiện biến đổi giá trị z bằng hàm Sigmoid có công thức bên dưới để đưa về khoảng xác suất:

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

Giá trị $\sigma(z) \in (0,1)$ được hiểu là xác suất mẫu thuộc về lớp dương (positive class). Đối với bài toán phân loại nhị phân (binary classification) thì mẫu sẽ được gán nhãn là 1 nếu xác suất lớn hơn 0.5 và được gán nhãn là 0 cho trường hợp ngược lại

Mã giả cho thuật toán LR đối với bài toán phân loại nhị phân:

Algorithm 1: Logistic Regression (Binary Classification)

Input: Training data
Output: Class labels
for $i = 1$ **to** k **do**
 for each training instance d_j **do**
 Compute predicted probability: $P(1|d_j)$;
 Set target: $z_j = \frac{y_j - P(1|d_j)}{P(1|d_j)(1 - P(1|d_j))}$;
 Set weight: $w_j = P(1|d_j)(1 - P(1|d_j))$;
 end
 Fit regression model $f(j)$ with targets z_j and weights w_j ;
end
Label decision: Assign class 1 if $P_{id} > 0.5$, else class 0;

Hình 3.3 Mã giả của mô hình Logistic Regression với bài toán phân loại nhị phân

Mở rộng cho phân loại đa lớp

Đối với bài toán phân loại nhiều hơn hai lớp như bài toán mà nhóm đang thực hiện (gồm có Positive, Neutral và Negative), LR được mở rộng bằng cách sử dụng phương pháp Softmax Regression – đây là phiên bản tổng quát của LR cho bài toán phân loại đa lớp

Softmax Regression tính xác suất một mẫu thuộc về từng lớp k trong tổng số K lớp như sau

$$P(y = k|x) = \frac{e^{w_k^T x}}{\sum_{j=1}^K e^{w_j^T x}}$$

Trong đó

x : vector đặc trưng của văn bản đầu vào

w_k : vector trọng số tương ứng với lớp k

K: Tổng số lớp trong bài toán

Hàm softmax đảm bảo tổng xác suất qua tất cả các lớp trên một mẫu là 1, và mô hình lựa chọn lớp có xác suất cao nhất để đưa ra nhãn dự đoán:

$$\hat{y} = \arg \max_k P(y = k|x)$$

Mã giả cho thuật toán LR đối với bài toán phân loại đa lớp:

Algorithm 2: Multinomial Logistic Regression (3-class classification)

Input: Training data $D = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$ with $y_i \in \{1, 2, 3\}$

Output: Weight matrix $W \in R^{3 \times d}$

Initialize weight matrix W randomly;

for $t = 1$ **to** T **do**

for each training sample (\mathbf{x}_i, y_i) **do**

 Compute scores: $z_j = \mathbf{w}_j^\top \mathbf{x}_i$ for $j = 1, 2, 3$;

 Compute softmax probabilities:

$$p_j = \frac{\exp(z_j)}{\sum_{k=1}^3 \exp(z_k)}$$

for $j = 1$ **to** 3 **do**

 Set indicator: $I_{ij} = \begin{cases} 1 & \text{if } y_i = j \\ 0 & \text{otherwise} \end{cases}$;

 Compute gradient: $\nabla_j = (p_j - I_{ij}) \cdot \mathbf{x}_i$;

 Update weights: $\mathbf{w}_j \leftarrow \mathbf{w}_j - \alpha \cdot \nabla_j$;

end

end

end

Hình 3.4 Mã giả của mô hình Logistic Regression cho bài toán phân loại đa lớp

Ưu điểm của LR:

- Dễ xây dựng, hiệu quả cho huấn luyện, ít tốn tài nguyên tính toán
- Dễ diễn giải bởi các trọng số được học từ các đặc trưng cho phép giải thích và trình bày rõ ràng về ảnh hưởng của từng từ đến xác suất phân lớp
- Hạn chế overfitting khá tốt với chính quy hóa (Regularization) như L1, L2, giúp tránh hiện tượng quá khớp khi làm việc với dữ liệu có số chiều lớn như các văn bản

Nhược điểm của LR:

- Vì LR giả định rằng dữ liệu có thể phân tách tuyến tính nên nó có thể không phù hợp với các bài toán mà ranh giới giữa các lớp khó phân biệt, chồng chéo lên nhau hoặc phi tuyến, ví dụ như các đánh giá 3 sao có thiên hướng khen nhiều hơn hoặc chê nhiều hơn
- Hiệu suất thường không cao bởi đặc trưng rời rạc được trích xuất từ các kỹ thuật trích xuất đặc trưng truyền thống như TF – IDF, mô hình chỉ có thể dựa trên tần suất xuất hiện của từng từ thay vì nhìn ra được ngữ cảnh của văn bản

- Khó có thể nhận ra được các mẫu cảm xúc theo dạng ngụ ý, mỉa mai hay đàm thoại vì LR không thể học các đặc trưng quá phức tạp

3.2.3 Linear Support Vector Machine - LinearSVM

LinearSVM là một mô hình phân loại tuyến tính thuộc một nhánh của mô hình nổi tiếng Support Vector Machine (SVM), với ý tưởng chính sử dụng đường siêu phẳng để phân tách các lớp sao cho biên phân cách (Margin) giữa các lớp là lớn nhất. Ở đồ án này, nhóm sử dụng LinearSVM để ứng dụng vào bài toán phân loại cảm xúc bằng cách kết hợp với kỹ thuật trích xuất đặc trưng là TF – IDF.

Trong các bài toán phân loại, LinearSVM cố gắng tìm một vector trọng số w và độ lệch b sao cho

$$f(x) = \text{sign}(w^T x + b)$$

Hàm sign ở đây đại diện cho hàm về dấu (sign function) được dùng để xác định nhãn phân loại đầu ra dựa trên giá trị của biểu thức tuyến tính $w^T x + b$

Các mô hình liên quan đến SVM phần lớn sẽ sử dụng hàm loss là **Hinge Loss**, mục tiêu của nó là đảm bảo hai tiêu chí, một là phân loại đúng, hai là phân loại đúng với biên an toàn (margin). Công thức của hàm như sau

$$L(y, f(x)) = \max(0, 1 - y \cdot f(x))$$

Trong đó

$y \in \{-1, 1\}$ là nhãn thật sự

$f(x) = w^T x + b$ là đầu ra của mô hình

Và mô hình sẽ giải quyết bài toán tối ưu sau

$$\min_{w,b} \left[\frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \max(0, 1 - y_i \cdot f(x)) \right]$$

Trong đó

w và b lần lượt là vector trọng số và độ lệch (bias)

C là siêu tham số điều chỉnh Regularization Strength

$\max(0.1 - y_i \cdot f(x))$ là hàm Hinge Loss

Sau khi huấn luyện, mô hình dự đoán các nhãn bằng cách tính \hat{y} như sau:

$$\hat{y} = \text{sign}(w^T x + b)$$

Mã giả của thuật toán

Algorithm 1: Linear Support Vector Classification (LinearSVC)

Input: Training data $D = \{(x_1, y_1), \dots, (x_N, y_N)\}$

Output: Decision function $f(x) = \text{sign}(w^\top x + b)$

1 Initialize weight vector $w = \mathbf{0}$, bias $b = 0$;

2 Minimize the following loss function:

3

$$\mathcal{L}(w, b) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \max(0, 1 - y_i(w^\top x_i + b))$$

4 **for** $t = 1$ **to** T (*max iterations*) **do**

5 | Compute sub-gradient of hinge loss w.r.t w, b ;

6 | Update parameters using an optimization method (e.g., SGD);

7 **foreach** *new input* x **do**

8 | Predict: $\hat{y} = \text{sign}(w^\top x + b)$;

Hình 3.5 Mã giả của mô hình LinearSVM

Ưu điểm:

- Mô hình có hiệu quả cao với dữ liệu có số chiều lớn như các văn bản
- Là mô hình tuyến tính nên vô cùng đơn giản và tốc độ thực thi cao
- Có thể trích xuất và giải thích được các đặc trưng quan trọng

Nhược điểm:

- Không xử lý tốt với dữ liệu phi tuyến tính
- Dễ nhạy cảm với nhiễu (noise) và giá trị ngoại lai (outliers) nếu không điều chỉnh tham số C một cách hợp lý
- Cần chuẩn hóa dữ liệu để đạt được hiệu quả tốt

3.3 Các độ đo đánh giá mô hình

3.3.1 Accuracy

Là một trong những độ đo cơ bản để đánh giá hiệu suất của các mô hình phân loại. Về định nghĩa, accuracy được tính bằng tổng số dự đoán đúng chia cho tổng số dự đoán. Công thức

$$\text{Accuracy} = \frac{\text{Tổng số dự đoán đúng}}{\text{Tổng số dự đoán}}$$

3.3.2 Precision

Là độ đo được sử dụng để đo lường tỷ lệ các mẫu được mô hình dự đoán đúng trong tổng số các mẫu mà mô hình đã dự đoán thuộc lớp đó, hiểu đơn giản ví dụ như với lớp Positive, Precision của lớp này sẽ được tính bằng số trường hợp dự đoán đúng là Positive chia cho tổng số trường hợp mà mô hình dự đoán là Positive. Dưới đây là công thức của độ đo Precision đối với lớp i cụ thể

$$\text{Precision}_i = \frac{TP_i}{TP_i + FP_i}$$

Trong đó:

- TP_i : Số lượng dự đoán đúng là lớp i
- FP_i : Số lượng mẫu mà mô hình dự đoán nhầm là lớp i

3.3.3 Recall

Là độ đo được sử dụng để đo lường tỷ lệ các mẫu thực sự thuộc lớp đó được mô hình dự đoán đúng. Hiểu đơn giản hơn thì với lớp Positive, Recall của lớp này sẽ được tính bằng số trường hợp dự đoán đúng là Positive chia cho tổng số lượng mẫu có nhãn thực tế là Positive. Dưới đây là công thức của độ đo Recall đối với lớp i cụ thể:

$$\text{Recall}_i = \frac{TP_i}{TP_i + FN_i}$$

Trong đó:

- FN_i : Số lượng mẫu thuộc lớp i bị dự đoán sai

3.3.4 F1 – Score

Là độ đo về trung bình điều hòa của Precision và Recall cho một lớp. Công thức cho lớp i:

$$F1\text{-Score}_i = \frac{2 \times \text{Precision}_i \times \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i}$$

3.3.5 Weighted F1 – Score

Là độ đo tổng hợp trong bài toán phân loại đa lớp, giúp đánh giá hiệu suất tổng thể của mô hình bằng cách tính trung bình giá trị F1-Score cho từng lớp nhưng điểm khác biệt là có gán trọng số về số lần xuất hiện của mỗi lớp trong tập dữ liệu

$$F1_{\text{weighted}} = \sum_{i=1}^K \frac{n_i}{N} \cdot F1_i$$

Trong đó

K là tổng số lớp

n_i là số mẫu thuộc lớp i

CHƯƠNG 4: THỰC NGHIỆM

4.1 Tiềm xử lý dữ liệu

Ở phần tiềm xử lý dữ liệu này, nhóm sẽ mô tả từng bước làm sạch, chuẩn hóa dữ liệu và giải thích lý do lựa chọn xử lý. Trong mỗi phần nhóm sẽ đưa ra các mẫu cụ thể trong tập dữ liệu cho từng nội dung xử lý

4.1.1 Xóa các mẫu bị trùng lặp và các mẫu không có đánh giá

Nhóm sử dụng hàm drop_duplicates của thư viện Pandas để thực hiện loại bỏ các mẫu bị trùng lặp và chỉ giữ lại mẫu đầu tiên với tham số keep='first' và hàm dropna để loại bỏ các mẫu không có nội dung đánh giá, các mẫu này thường sẽ là các đánh giá chỉ chấm điểm trên thang đo sao mà không để lại bình luận hay góp ý, không có giá trị cho việc phân tích. Sau khi thực hiện bước này thì tập dữ liệu huấn luyện giảm từ 90000 mẫu còn 85076 mẫu và tập dữ liệu kiểm tra giảm từ 10000 mẫu còn 9738 mẫu

4.1.2 Trích xuất và loại bỏ

4.1.2.1 Các phần chưa được xử lý liên quan đến HTML

Bộ dữ liệu này theo như nhóm quan sát, phân tích thì được thực hiện từ quá trình cào dữ liệu trên trang web của Amazon và các đánh giá này vẫn còn sót lại những phần chưa xử lý từ HTML, chẳng hạn các tag như
 hay các tokens đặc biệt gồm cả số và chữ như '", ví dụ như các đánh giá bên dưới. Vì vậy nhóm thực hiện loại bỏ các pattern vì chúng không mang lại ý nghĩa trong phân tích

I wanted a pair of nice pool shoes for a trip to Europe, and though I've never owned 'Jacks,' these were cute online and solid black, so I decided to give them a try.

The only other Jack Rogers I own (I bought them at the same time) are the metallic wedges, so I can't compare the fit to the other flat sandals, but these fit really loosely. The length doesn't seem to run big, it's how much room there is between the sole and the top of the straps. It doesn't fit snugly against my foot, so it is constant effort to keep them on. I do have a narrower foot, but I've never had that happen with sandals.

My other complaint about this (especially since I bought them with the intention of being around the pool and water) is that these are incredibly slippery. Like, don't walk down stairs slippery. There is no grip, to the point that I fell down our wood stairs because my foot slid out from under me.

/>These are really cute, and I wanted to like them (I bought the solid black with the little hold stud accents), but they are so poorly fitted and slippery that I'm just going to list them on eBay for a loss. Totally impractical.

the "crystals" on the top of the thong look like cheap plastic, since that is what they are. I liked the fabric, but could not get off the fake crystals without damaging the shoe. I returned them

4.1.2.2 Các tokens vừa chứa số vừa chứa kí tự

Trong các đánh giá, sẽ có những mẫu mà khách hàng sẽ gửi đính kèm các video, có thể dùng để ghi lại quá trình mở và kiểm hàng cũng như review sơ lược về sản phẩm, hoặc các mã sản phẩm, đây cũng là thông tin không cần thiết, vì vậy nhóm quyết định xóa các pattern này

[[VIDEOID:1051f3a729a3ee387754a87d00513763]]Wanted to see what the review looks like when I use "quotes" because there may be an encoding thing with opera? Hmm everything looked ok, *trying* some other 'things' to see if they are problematic.

After reading about all the fakes out there, I decided to purchase directly from Ray-Ban online. I also wanted red with polarized lenses, which they didn't make unless I paid extra for customization. I had tried on many pairs of the RB2132s at Sunglass Hut, in all sorts of colors of frames and lenses, polarized and not, so I knew exactly what I wanted. So disappointed with what came. If I hadn't bought directly from Ray-ban I would have been sure they were fake. They were cheap plastic, and had no metal supports down the temples. They squeak when opening the temples or putting them on. They will no doubt break in short order. Now I see the other reviews that talk about different manufacturing than in the past. Oh, how I wish I would have read these first.

4.1.2.3 Các email, website hoặc số điện thoại

Đây cũng là các thông tin không mang lại ý nghĩa cho việc phân tích cảm xúc nên nhóm cũng thực hiện loại bỏ

These shoes are grippy and look good. They are comfortable enough during break in. HOWEVER, I never really got to break them in without the back tearing. I've been on only two mtb rides with these. Now I'm sitting in an airport looking like a fool with a ripped shoe. 510, I'd like my money back - trevordunn87@gmail.com

Some Ray Ban models are produced in China but this particular model doesn't appear to be one of them. See the link below for some helpful information:
<http://www.ebay.com/gds/Are-authentic-Ray-Bans-really-being-made-in-China-/10000000020632354/g.html>

4.1.3 Chuẩn hóa và ánh xạ

4.1.3.1 Chuẩn hóa các dấu nháy

Để đồng nhất các mẫu, nhóm quyết định chuẩn hóa các dấu nháy như ' , ` , ' về cùng một định dạng nháy là ' . Việc chuẩn hóa như vậy nhằm mục đích phục vụ cho các bước xử lý sau để tách các từ viết tắt thành các từ riêng biệt như don't trở thành do not

4.1.3.2 Chuẩn hóa các giá trị phân số về thập phân

Trong các đánh giá của khách hàng, đối với các size giày, họ thường ghi bằng phân số, ví dụ như 8 ½, theo ý kiến của nhóm thì nếu để nguyên sẽ gây khó cho mô hình khi học các đặc trưng, vì vậy nhóm quyết định chuyển hóa các số này về số thập phân, thay vì ½ sẽ là 8.5, hàm chuẩn hóa sẽ dành riêng cho số ½ vì đối với chủ đề về giày dép, ngoài phân số này ra thì các phân số còn lại hiếm khi và gần như không xuất hiện

4.1.3.3 Chuyển đổi và hiển thị các từ ngữ thô tục

Trong các đánh giá sẽ có vài khách hàng chèn thêm các từ ngữ thô tục để biểu đạt cảm xúc của mình, tuy nhiên các từ này thường sẽ bị che đi bằng các kí tự * hoặc viết theo một cách khác để không hiển thị quá rõ ràng, tránh bị kiểm duyệt hay bị xóa bỏ. Nhóm nhận thấy rằng các từ này rất đáng để khai thác và cho mô hình học, vì vậy nhóm quyết định thực hiện chuyển hóa về dạng đầy đủ và hiển thị rõ ràng từ như “bitch”, “fuck”, “ass”,.. thay vì bị che hoặc bị viết theo kiểu lách

I wanted a pair of black pointed toe flats so I bought the Zariner and the Itzamor loafers to compare.
Sizing:
I'm a size 6 and both shoes felt true to size.
Comfort:
I have wide feet and both the Zariner and Itzamor felt a bit too narrow for me. These would be fine with people with narrow feet but people with 'normal' or wide feet would find them too tight in the toe area. I do not think I would have been able to walk very far with these shoes on. I felt a pinch between my pinky toe and my 4th toe whenever I walked in either shoe. The bottom of both shoes are also a little slippery. If either loafers were comfortable for me, I

would have taken them to a shoe cobbler to add a non-slip sole so I wouldn't slip and bust my @ss.

Style:
I like the style of the Zariner. It's got a plain toe but has a thin gold trim on the heel for some subtle flashiness. The leather was a little softer than the Itzamor.
The Itzamor has a fabric tweed-like upper which I did not love because it would get soaked if worn in the rain. I liked the patent leather but it felt a bit stiff.

Overall, these shoes look cute but did not feel comfortable enough to walk in. They were returned.

4.1.3.4 Chuyển đổi các biểu tượng cảm xúc

Nhóm nhận thấy rằng có những đánh giá không sử dụng từ ngữ để đánh giá mà chỉ sử dụng các biểu tượng cảm xúc (emoji) để thể hiện cảm xúc của khách hàng đối với các sản phẩm, ví dụ như các đánh giá cho lớp Positive sẽ gồm các emoji như thả dấu “like”, “mặt cười”, “100 điểm”, “trái tim”,... còn các đánh giá cho lớp Negative phần lớn là dấu “like” nhưng quay ngược lại hoặc các emoji theo dạng buồn, tiếc nuối. Các emoji này sẽ có hai hướng xử lý, một là loại bỏ, hai là sẽ ánh xạ sang các tính từ tương ứng, nhóm cho rằng việc ánh xạ sang các tính từ là hợp lý hơn, vì việc chuyển đổi emoji sang tính từ không chỉ bảo toàn sắc thái cảm xúc trong dữ liệu mà còn tăng cường số lượng từ ngữ đặc trưng cho từng lớp, hỗ trợ mô hình khai thác tốt hơn

Emoji	Tính từ ánh xạ
👍	Satisfied
😊	Happy
😍	Delighted
💯	Excellent
👎	Disappointed
😔	Sad
😢	Discouraged

Phần lớn thì các emoji này đại diện cho các hai lớp Positive và Negative, và các đánh giá ở lớp Neutral hiếm xuất hiện các đánh giá chỉ toàn emoji, nên việc chuyển đổi này sẽ có lợi cho hai lớp là Positive và Negative

4.1.3.5 Chuyển đổi các từ lóng

Các từ lóng (slang) hoặc các từ viết tắt thường được mọi người sử dụng, kể cả trong tiếng Anh và tiếng Việt, giúp giảm thời gian viết, tuy nhiên các từ này thường không nằm trong từ điển thông thường mà do mọi người tự quy định với nhau, và chúng thường sẽ khiến cho các mô hình khó hiểu, khó học được cũng như nắm bắt được ngữ cảnh chính xác. Vì vậy nhóm thực hiện trích xuất các từ lóng và từ viết tắt trong bộ dữ liệu bằng cách kiểm tra với từ điển trong thư viện NLTK, trong quá trình trích xuất thì vẫn xuất hiện các từ có nghĩa thông thường nhưng bị liệt kê nằm ngoài từ điển, lý do là vì các từ này thường là động từ được chia theo thì, dạng bị động hoặc dạng rút gọn. Do đó, nhóm phải kết hợp giữa kiểm tra tự động và thực hiện liệt kê thủ công các từ lóng và từ viết tắt. Dưới đây là bảng tóm tắt các từ lóng và từ viết tắt, vì số lượng khá lớn nên nhóm chỉ chọn ra khoảng 4-5 từ để đưa vào bảng

Các từ lóng, từ viết tắt	Các từ chuẩn thông thường
omg	oh my god
lol	laughing out loud
fav	favorite
btw	by the way
sz	size

4.1.3.6 Chuyển đổi các mẫu bị lặp chữ cái trong các từ

Nhóm nhận thấy rằng có một số đánh giá mà người dùng thực hiện việc “spam” các kí tự trong một từ, điều này có thể là khách hàng thể hiện cảm xúc mạnh như quá vui hoặc quá tức giận, ví dụ như token hhhaaaappppyyyyyy được sử dụng để nhấn mạnh từ cụm từ “happy”. Nếu như giữ nguyên mà không xử lý thì các từ này sẽ không mang lại ý nghĩa quá lớn, nếu số lượng lặp lại cho từng kí tự là như nhau hoặc số lượng lặp lại cho cùng một token là giống nhau (cùng một quy tắc lặp) thì sẽ mang lại nhiều ý nghĩa, nhưng đối với trường hợp này thì họ sẽ spam một cách ngẫu nhiên các từ. Vì vậy nhóm thực hiện việc chuẩn hóa theo quy tắc là nếu số kí tự liên tiếp lặp lại từ 3 trở lên thì sẽ rút gọn thành 1, nhóm không lựa chọn từ 2 trở lên vì sẽ có những từ có 2 kí tự liên tiếp nhưng vẫn đúng, ví dụ như từ “too” hay “feel”

Văn bản gốc	Văn bản sau khi chỉnh sửa
what are ttttthhhhhhioooooosssssseeeeeeee	what are thiose

4.1.4 Sửa lỗi chính tả

Sau khi hoàn tất các bước tiền xử lý trước đó, nhóm tiếp tục thực hiện kiểm tra và sửa lỗi chính tả nhằm đảm bảo rằng các token trong văn bản đều mang ý nghĩa rõ ràng và chính xác. Để thực hiện bước này, nhóm sử dụng thư viện **TextBlob**, một công cụ hỗ trợ phát hiện và tự động hiệu chỉnh các từ viết sai chính tả. Việc sửa lỗi chính tả không chỉ giúp nâng cao chất lượng biểu diễn đặc trưng mà còn giúp mô hình học máy tiếp cận dữ liệu đầu vào một cách chính xác và hiệu quả hơn.

Văn bản sau khi thực hiện các bước trước	Văn bản sau khi được sửa lỗi chính tả
what are thiose	what are those

4.1.4.1 Tách các từ viết dính liền bằng phương pháp thủ công

Nhóm thực hiện truy xuất các mẫu có các token có độ dài bất thường với ngưỡng đặt ra là 15 và 18, vì ít có từ ngữ nào có độ dài lớn hơn các giá trị nói trên, sau khi trích xuất và kiểm tra thì nhóm nhận thấy có các trường hợp thực sự là các từ có độ dài bất thường này là từ có nghĩa, phần lớn là các tính từ, bên cạnh đó cũng có các mẫu chứa các từ dính liền với nhau và một số từ viết sai chính tả, nhóm thực hiện truy xuất index của mẫu và ánh xạ thủ công, dưới đây là một số trường hợp viết dính liền mà nhóm nhận thấy

Các từ bị viết dính liền	Các từ được phân tách
getmyfactsstraight	get my facts straight
supercute	super cute
completlydifferent	completely different
itemswerenotshoes	items were not shoes
quitecomfortable	quite comfortable

4.1.4.2 Loại bỏ các stopwords

Ở bước cuối cùng, để phục vụ cho trích xuất đặc trưng từ TF – IDF, nhóm thực hiện loại bỏ các stopwords để tránh cho các mô hình học các đặc trưng không mang lại ý nghĩa nhiều, các stopwords ở đây gồm các đại từ nhân xưng, đại từ sở hữu, mạo từ, giới từ,... Tuy nhiên nhóm thực hiện tạo thêm một cột trong dữ liệu để lưu các mẫu sau khi loại bỏ stopwords và vẫn giữ lại các mẫu trước khi loại bỏ để phục vụ cho việc phân tích và giải thích kết quả

Sau khi hoàn tất các bước xử lý dữ liệu trên, nhóm đánh giá rằng ngũ liệu đã được làm sạch khoảng 80%, do vẫn còn tồn tại một số vấn đề liên quan đến ngũ pháp và ngũ nghĩa. Điều này là dễ hiểu vì văn phong trong các đánh giá của người dùng thường mang tính tự do, thiếu chuẩn mực như các văn bản hành chính, dẫn đến việc khó xử lý triệt để bằng các quy tắc ngôn ngữ học truyền thống.

4.2 Huấn luyện và kiểm thử các mô hình

Như phần phương pháp đã đề cập, nhóm thực hiện huấn luyện và so sánh ba mô hình máy học kết hợp với kỹ thuật trích xuất đặc trưng TF – IDF. Đối với mỗi mô hình, nhóm thực hiện tìm tham số tối ưu với hai kỹ thuật phổ biến là RandomizedSearchCV và GridSearchCV. Đầu tiên nhóm sử dụng RandomizedSearchCV với các giá trị của từng tham số có độ lệch lớn để thu hẹp phạm vi tìm kiếm, sau khi có được tham số từ RandomizedSearchCV, nhóm tiếp tục thực hiện tìm kiếm các tham số tối ưu bằng GridSearchCV với các giá trị lân cận giá trị tham số nhận được từ RandomizedSearchCV. Việc kết hợp hai bước này giúp giảm chi phí tính toán so với việc dùng GridSearchCV một cách trực tiếp trên không gian tham số rộng, đồng thời đảm bảo hiệu quả trong việc tìm kiếm cấu hình tốt nhất cho từng mô hình.

4.2.1 MNB + TF – IDF

Sau khi thực hiện tìm kiếm tham số tối ưu thì kết quả đạt được như sau

Bảng 4.1 Giá trị tham số tối ưu của MNB và TF-IDF

Tham số	Giá trị thử nghiệm	Giá trị tối ưu
max_features	10000, 11000, 12000	12000
n_gram_range	(1, 1), (1, 2), (1, 3)	(1, 3)
min_df	3, 4, 5	4
sublinear_df	True, False	False
use_idf	True, False	False

Và bảng classification report gồm các độ đo như Accuracy, Precision, Recall, F1-Score, Weight-F1 của mô hình MNB như sau:

Bảng 4.2 Classification Report của mô hình MNB

	Precision	Recall	F1-Score
Negative	0.77	0.74	0.76

Neutral	0.41	0.43	0.42
Positive	0.79	0.80	0.80
Weighted Avg	0.71	0.70	0.70
Accuracy	0.70		

4.2.2 $LR + TF - IDF$

Tham số tối ưu của mô hình LR kết hợp với TF – IDF như sau

Bảng 4.3 Giá trị tham số tối ưu của mô hình LR và TF-IDF

Tham số	Giá trị thử nghiệm	Giá trị tối ưu
max_features	5000, 10000, 15000	15000
n_gram_range	(1, 1), (1, 2), (1, 3)	(1, 2)
min_df	1, 2, 3	2
sublinear_df	True, False	True
use_idf	True, False	False
norm	'l1', 'l2'	'l2'
C	0.5, 1, 1.5	1.5

Bảng classification report của mô hình LR như sau:

Bảng 4.4 Classification Report của mô hình LR

	Precision	Recall	F1-Score
Negative	0.75	0.86	0.80
Neutral	0.46	0.28	0.35
Positive	0.80	0.85	0.83
Weighted Avg	0.71	0.74	0.72
Accuracy	0.74		

4.2.3 $SVM + TF - IDF$

Tham số tối ưu của mô hình SVM

Bảng 4.5 Giá trị tham số tối ưu của mô hình SVM

Tham số	Giá trị thử nghiệm	Giá trị tối ưu
max_features	5000, 10000, 15000	15000
n_gram_range	(1, 1), (1, 2), (1, 3)	(1, 2)
min_df	1, 2, 3	3

sublinear_tf	True, False	True
use_idf	True, False	False
norm	‘l1’, ‘l2’	‘l2’
C	0.3, 0.5, 0.7	0.5
loss	‘hinge’, ‘squared_hinge’	‘squared_hinge’

Classification Report của mô hình SVM

Bảng 4.6 Classification Report của mô hình SVM

	Precision	Recall	F1-Score
Negative	0.74	0.87	0.80
Neutral	0.46	0.22	0.30
Positive	0.79	0.85	0.82
Weighted Avg	0.71	0.74	0.71
Accuracy		0.74	

Bảng tổng kết về độ đo accuracy và f1-weighted của ba mô hình khi kết hợp với trích xuất đặc trưng TF – IDF với tham số tối ưu

Bảng 4.7 Tóm tắt kết quả của ba mô hình

Các mô hình	Accuracy	F1-Weighted
MNB	0.70	0.70
LR	0.74	0.72
SVM	0.74	0.71

CHƯƠNG 5: NHẬN XÉT VÀ PHÂN TÍCH KẾT QUẢ

Để thực hiện việc phân tích và nhận xét về kết quả đạt được của từng mô hình, nhóm sử dụng hai bước chính là trích xuất các hệ số đặc trưng của từng mô hình và in ra một vài dự đoán sai của các mô hình để phân tích

5.1 Multinomial Naïve Bayes

Đối với mô hình MNB, nhóm sử dụng thuộc tính `feature_log_prob_`, thuộc tính này biểu thị cho giá trị của log xác suất có điều kiện của từng đặc trưng ứng với từng lớp nhãn. Về ý nghĩa, các giá trị này cho biết mức độ đóng góp của từng từ vựng vào xác suất dự đoán của mô hình, từ đó giúp quan sát và nhận diện được các từ đặc trưng nổi bật cho từng cảm xúc. Công thức của `feature_log_prob_` vô cùng đơn giản, đầu tiên nó sẽ lưu trữ lại logarithm của xác suất có điều kiện như sau

$$\log P(x_j | y_k)$$

Trong đó x_j là đặc trưng, cụ thể hơn là một từ trong từ vựng, y_k là lớp nhãn thứ k .

Xác suất $P(x_j | y_k)$ được tính dựa trên Smoothing Laplace là

$$P(x_j | y_k) = \frac{N_{kj} + \alpha}{\sum_{l=1}^V (N_{kl} + \alpha)} = \frac{N_{kj} + \alpha}{N_k + \alpha \cdot V}$$

Từ đó `feature_log_prob_` của lớp k với đặc trưng j được tính như sau

$$\text{feature_log_prob_[k, j]} = \log \frac{N_{kj} + \alpha}{N_k + \alpha \cdot V}$$

Đối với mỗi lớp, sẽ có một danh sách các đặc trưng đi kèm với giá trị `feature_log_prob_` tương ứng, giá trị càng lớn thì đặc trưng đó càng mang ý nghĩa. Danh sách và giá trị các đặc trưng được đánh giá và xếp hạng về mức độ ý nghĩa sau quá trình huấn luyện

Bảng 5.1 Giá trị về các hệ số đặc trưng của từng lớp trong mô hình MNB

Negative		Neutral		Positive	
Đặc trưng	Giá trị tương ứng	Đặc trưng	Giá trị tương ứng	Đặc trưng	Giá trị tương ứng
not	-2.9887	not	-3.0474	not	-3.3171
feet	-4.5472	feet	-4.4131	feet	-4.3060

shoe	-4.6484	size	-4.4895	very	-4.4991
really	-4.6635	shoe	-4.6506	great	-4.5433
size	-4.6640	really	-4.6681	comfortable	-4.5826
too	-4.6844	too	-4.7504	so	-4.6123
very	-4.7097	so	-4.7974	love	-4.6333
so	-4.8240	very	-4.8232	really	-4.7005
wear	-5.1745	wear	-4.9627	size	-4.7041
not not	-5.1957	not not	-5.2668	wear	-4.8119

Thông qua các giá trị nêu trên, nhóm có một vài nhận xét và phân tích về mô hình MNB kết hợp với TF-IDF như sau:

- Ở cả ba nhãn, các từ đơn phổ biến như “not”, “feet”, “size”, “really”, “very”, “wear” đều xuất hiện. Điều này phản ánh rằng các từ này có tần suất xuất hiện cao trên bộ dữ liệu huấn luyện, nhưng không mang lại nhiều ý nghĩa giúp phân biệt giữa các lớp một cách rõ ràng
- Mặc dù tham số tối ưu của TF-IDF là n_gram_range khi hoạt động với MNB là (1, 3) nhưng phần lớn các từ lại là từ đơn, chỉ có một trường hợp là “not not”, tuy nhiên cụm từ này có mức độ ảnh hưởng không quá lớn đến mô hình khi mà giá trị log khá thấp, theo như nhóm quan sát thì từ này xuất hiện do cách đánh giá của khách hàng, họ nhấn mạnh về sự không hài lòng đối với sản phẩm
- Ở hai lớp Neutral và Negative, các đặc trưng gần như tương đồng với nhau, thiếu đi các đặc trưng cảm xúc rõ nét, gây khó khăn cho mô hình trong việc dự đoán. Tuy nhiên đối với lớp Positive vẫn tồn tại các từ mang cảm xúc rõ ràng như “great”, “comfortable”, “love”

Bảng 5.2 Một số trường hợp MNB dự đoán nhãn khác với nhãn thực tế

Mẫu trước khi loại bỏ stopwords	Mẫu sau khi loại bỏ stopwords	Nhãn thực tế	Nhãn dự đoán	Giải thích kết quả
they are are not comfortable what so ever and i got	not comfortable so blisters so no longer wear	Neutral	Negative	Đánh giá của khách hàng nằm ở mức 3 sao nên được phân thành Neutral, tuy

blisters so i no longer wear them				nhiên sắc thái đánh này mang tính chê bai và tiêu cực, nên đây là một dự đoán hợp lý của mô hình
delivery and everything was fine however but i cannot put anything with much weight in it is without the candles coming out of place	delivery everything fine however cannot put anything much weight without candles coming place	Negative	Positive	Đánh giá này vừa có khen vừa có chê, tuy nhiên sắc thái nghiêng về chê nhiều hơn, mô hình dự đoán sai, nguyên nhân có thể là do từ “fine” có ảnh hưởng lớn trong quá trình huấn luyện của lớp Positive
i love the shoe very comfortable feet it is perfect then the second time i wore them i leaned back and the heel popped i hope i cannot fix them this was the shoe that is i was going to try in all the colors but probably not now	love shoe very comfortable feet perfect then second time wore leaned back heel popped hope cannot fix shoe going try all colors probably not now	Neutral	Positive	Giống như đánh giá trên thì ở đánh giá này vừa có khen vừa có chê, tuy nhiên xuất hiện đồng thời hai đặc trưng có mức độ ảnh hưởng lớn đến lớp Positive là love và comfortable nên mô hình dự đoán sai khi cho rằng đây là đánh giá mang thiên hướng tích cực
they are worked well enough for a	worked well enough wedding	Neutral	Positive	Ở đánh giá này thì khách hàng vừa có

wedding i love the look but had to take them off under the table at dinner because they are we are hurting my feet	love look take off table dinner because hurting feet			khen vừa có chê, tuy nhiên các đặc trưng thiên về lớp Positive lại chiếm ưu thế hơn so với Negative khi có “well”, “love” so với “hurting” nên mô hình dự đoán nghiêng về Positive hơn
very good	very good	Negative	Positive	Đây chắc hẳn là đánh giá nhầm trên thang sao của khách hàng bởi đánh giá này chắc chắn là tích cực, tuy nhiên họ có thể đánh nhầm sang 1 hoặc 2 sao, mặc dù vậy thì mô hình vẫn dự đoán được chính xác

5.2 Logistic Regression

Đối với mô hình Logistic Regression, mô hình học một vector trọng số là w sao cho dự đoán xác suất của mỗi lớp được tối ưu hóa qua hàm softmax cho bài toán nhiều lớp. Với mỗi đặc trưng x_j , mô hình học được một hệ số w_j phản ảnh mức độ ảnh hưởng của đặc trưng đó đến khả năng phân loại vào một nhãn nhất định nào đó. Công thức xác suất cho mỗi lớp với một vector trọng số riêng w_k như sau

$$P(y = k | x) = \frac{e^{w_k^T x}}{\sum_j e^{w_j^T x}}$$

Và các trọng số w_k được học thông qua hàm mất mát log-loss kết hợp với các regularization.

Về ý nghĩa, nếu $w_j > 0$ thì đặc trưng x_j tăng xác suất thuộc lớp tương ứng và ngược lại. Đối với $|w_j|$, giá trị này càng lớn thì đặc trưng càng có sức ảnh hưởng mạnh đến dự đoán của mô hình. Vì vậy nhóm sử dụng thuộc tính `coef_` để truy xuất các đặc trưng quan trọng nhất trong từng lớp qua quá trình huấn luyện của Logistic Regression và có được bảng thống kê như sau

Bảng 5.3 Giá trị về các hệ số đặc trưng của từng lớp trong mô hình LR

Negative		Neutral		Positive	
Đặc trưng	Giá trị tương ứng	Đặc trưng	Giá trị tương ứng	Đặc trưng	Giá trị tương ứng
horrible	4.4802	not most	2.6687	not hurt	4.2412
terrible	4.2496	okay not	2.1954	love	4.1885
poor	4.1250	not bad	2.1027	perfect	4.0229
not happy	4.0353	okay	2.0739	great	3.8914
not money	4.0273	not best	1.8961	loves	3.8192
worst	4.0243	broke month	1.8543	favorite	3.6081
not worth	3.7957	than most	1.8071	not disappointed	3.5831
no good	3.7576	cannot really	1.8059	compliments	3.5453
disappointed	3.7246	higher	1.8028	best	3.5388
wanted love	3.4994	max	1.7464	perfectly	3.5361

Qua danh sách về các hệ số đặc trưng của mô hình Logistic Regression kết hợp với TF - IDF, nhóm có một vài nhận xét sau

- So với bảng hệ số đặc trưng từ mô hình Multinomial Naive Bayes, mô hình Logistic Regression thể hiện sự phân hoá rõ ràng hơn giữa các lớp. Các đặc trưng trọng yếu trong từng lớp ít trùng lặp hơn, cho thấy mô hình có khả năng học được những dấu hiệu phân biệt tốt hơn giữa các nhãn.

- Với lớp Negative, các đặc trưng như “horrible”, “terrible”, “worst”, “poor” đều là các từ có sắc thái tiêu cực rõ rệt. Tuy nhiên, một số cụm như “wanted love” có thể không mang sắc thái tiêu cực trực tiếp, điều này phản ánh những mâu ngứ nghĩa phức tạp hoặc phụ thuộc ngữ cảnh mà mô hình vẫn coi là có giá trị phân loại.
- Với lớp Neutral, các cụm từ như “not most”, “okay not”, “not bad”, “not best”, hay “cannot really” thường là các biểu thức mang tính trung lập hoặc lưỡng lự, thường xuất hiện trong các đoạn đánh giá không quá thiên về tích cực hay tiêu cực. Đây là những biểu thức thường được dùng khi khách hàng thuật lại trải nghiệm một cách trung lập, không đưa ra đánh giá rõ ràng.
- Đối với lớp Positive, các đặc trưng như “love”, “perfect”, “favorite”, “best”, “compliments” mang sắc thái tích cực rõ ràng. Đặc biệt, các cụm như “not hurt”, “not disappointed” tuy mang từ phủ định nhưng ngữ nghĩa lại tích cực, cho thấy mô hình Logistic Regression có khả năng nhận diện được mâu từ gián tiếp (như phủ định của điều tiêu cực).
- Sự tách biệt rõ ràng về đặc trưng của ba lớp là lý do mà độ hiệu quả và ổn định của LR tốt hơn so với MNB về các độ đo nhờ vào khả năng gán trọng số linh hoạt và không quá phụ thuộc vào xác suất như MNB đã làm

Bảng 5.4 Một số trường hợp LR dự đoán nhãn khác với nhãn thực tế

Mẫu trước khi loại bỏ stopwords	Mẫu sau khi loại bỏ stopwords	Nhãn thực tế	Nhãn dự đoán	Giải thích kết quả
i like how they are look but they are are too small for my face	how look too small face	Neutral	Negative	Việc loại bỏ các stopword không hợp lý ảnh hưởng khá nhiều đến việc dự đoán bởi đây là một đánh giá rõ ràng vừa có khen vừa có chê nhưng việc loại bỏ

				stopwords vô tình làm mắt luôn vẻ khen làm cho mô hình dự đoán sai
looked fine too small returned	looked fine too small returned	Neutral	Negative	Đánh giá này vừa có khen vừa có chê, tuy nhiên sắc thái tiêu cực chiếm phần ưu khi vừa có cụm “small” và “returned”, điều này cho thấy mô hình dự đoán hợp lý
they are run about half size small	run half size small	Neutral	Negative	Đánh giá mang tính thông báo nhưng từ small là đặc trưng ảnh hưởng lớn của lớp Negative
too small lenses too close together and not wide enough	too small lenses too close together not wide enough	Neutral	Negative	Đánh giá thiên nhiều về tiêu cực khi phần lớn là chê, đây là đánh giá 3 sao nhưng có phần chê nhiều hơn, và mô hình dự đoán hợp lý
very good	very good	Negative	Positive	Gặp lỗi tương tự như đã phân tích ở MNB

5.3 SVM

Sau khi huấn luyện, SVM lưu các trọng số w trong thuộc tính `coef_`. Mỗi hàng trong `coef_` tương ứng với một lớp và mỗi cột tương ứng với một đặc trưng. Cũng như LR, hệ số dương thì đặc trưng góp phần tăng mẫu đang xét thuộc về một lớp cụ thể và ngược lại, cũng như với tri tuyệt đối của hệ số thì giá trị này càng lớn thì đặc trưng đó càng ảnh hưởng lớn đến việc đưa ra dự đoán của mô hình cho lớp tương ứng. Khác với MNB thì `coef_` là trọng số tuyến tính mà không phải xác suất, mô hình sẽ lấy tổng

Bảng thống kê về hệ số đặc trưng của mô hình LinearSVM kết hợp TF – IDF như sau

Bảng 5.5 Giá trị về các hệ số đặc trưng của từng lớp trong mô hình SVM

Negative		Neutral		Positive	
Đặc trưng	Giá trị tương ứng	Đặc trưng	Giá trị tương ứng	Đặc trưng	Giá trị tương ứng
not happy	2.5754	little disappointed	2.1273	not hurt	4.2412
worst	2.5582	not most	1.9852	love	4.1885
no good	2.5407	okay not	1.7281	perfect	4.0229
horrible	2.4250	top very	1.6129	great	3.8914
most uncomfortable	2.4240	sure great	1.5767	loves	3.8192
terrible	2.4209	not bad	1.5716	favorite	3.6081
useless	2.3836	max	1.5479	not disappointed	3.5831
wanted love	2.3727	beautiful not	1.5442	compliments	3.5453
not money	2.3262	not cushion	1.5123	best	3.5388
not worth	2.2765	wear quickly	1.4947	perfectly	3.5361

Một vài nhận xét và đánh giá của nhóm từ các hệ số đặc trưng:

- SVM cũng làm khá tốt so với MNB khi thể hiện sự phân hóa một cách rõ ràng giữa các đặc trưng, các đặc trưng trong mỗi lớp hầu như chỉ xuất hiện một lần, không bị trùng lặp giữa các lớp

- Với lớp Negative, vẫn là các từ mang rõ sắc thái tiêu cực như “horrible”, “terrible”, “useless”, các cụm như “no money”, “no worth” có thể là đánh giá mà người dùng cho rằng sản phẩm không phù hợp với túi tiền hoặc giá trị mang lại
- Với lớp Neutral, ngoài các cụm mang sắc thái trung tính thì vẫn có các cụm liên quan đến mô tả trải nghiệm thực tế của khách hàng như “wear quickly”, “not cushion”
- Với lớp Positive vẫn có các từ đặc trưng mạnh cho việc khen ngợi, đặc biệt phần lớn sẽ là kết hợp giữa từ “not” và các tính từ tiêu cực, các từ như “not hurt”, “perfectly” cho thấy những trải nghiệm tốt từ khách hàng
- So với LR thì các đặc trưng của SVM có phần trùng lặp, tuy nhiên hệ số từ SVM thấp hơn, nguyên nhân có thể là do SVM tập trung vào sự phân tách biên giữa các lớp hơn là xác suất xuất hiện

Bảng 5.6 Một số trường hợp SVM dự đoán nhãn khác với nhãn thực tế

Mẫu trước khi loại bỏ stopwords	Mẫu sau khi loại bỏ stopwords	Nhãn thực tế	Nhãn dự đoán	Giải thích kết quả
she is look amazing but run smaller than i was expecting also the sole of one of the she is was coming off i had to superglue it is not a problem but should not have not had to do that is with new she is	look amazing run smaller than expecting sole coming off superglue not problem not not new	Neutral	Negative	Các từ như smaller, problem, not not làm cho mô hình dự đoán nghiêng về nhãn Negative mặc dù đánh giá này mang tính trung lập
i had 3 friends listen while i wore them and traveled at different speed	3 friends listen wore traveled different speed said same thing	Positive	Negative	Đánh giá này mang sắc thái tích cực nhưng không quá rõ ràng thông qua

and they are said the same thing				các từ ngữ, kể về trải nghiệm tốt cả khách hàng và bạn bè của họ, tuy nhiên ngữ nghĩa không rõ ràng cùng với từ different có thể là nguyên nhân khiến cho mô hình dự đoán thành Negative
this was my second pair of sank s i was disappointed in this pair i cannot only wear about an 1 hour due to hurting the back of my heal	second pair sank s disappointed pair cannot only wear 1 hour due hurting back heal	Neutral	Negative	Đánh giá mặc dù nằm ở mức 3 sao nhưng sắc thái chủ đạo là tiêu cực bởi trải nghiệm không tốt từ khách hàng thông qua các cụm như disappointed, hurting và mô hình vẫn dự đoán hợp lý
i ordered my size 8 5 b but these she is we are to wide on my foot i liked the looks of them but they are we are just to big and i sent them back	ordered size 8 5 b wide foot liked looks big sent back	Neutral	Negative	Cũng như đánh giá trên, đây là đánh giá thiên về tiêu cực nhiều hơn bởi kích thước không vừa và khiến họ phải trả lại hàng, dự đoán của mô hình là chấp nhận được

i like the shoe just fine but the faces are annoying they are thin and i do not know what to do with the big loop that is left	shoe fine faces annoying thin not know big loop left	Neutral	Negative	Cũng như đánh giá trên, về nghĩa gốc thì đánh giá này hoàn toàn mang sắc thái tiêu cực, đây là đánh giá 3 sao nhưng trải nghiệm và thái độ của họ về sản phẩm đều là không tốt, đây là dự đoán hợp lý của mô hình
--	--	---------	----------	---

CHƯƠNG 6: KẾT LUẬN

Trong đồ án này, nhóm đã thực hiện thử nghiệm ba mô hình học máy là Multinomial Naïve Bayes, Logistic Regression và SVM, mỗi mô hình kết hợp với kĩ thuật trích xuất đặc trưng là TF – IDF để thực hiện nhận diện cảm xúc trên các đánh giá sản phẩm trên sàn thương mại điện tử Amazon.

Từ kết quả về các độ đo của ba mô hình, nhóm cho rằng các mô hình học máy đơn giản có thể nhận diện và phân loại cảm xúc ở mức độ khá vì các độ đo và hiệu suất chỉ nằm trong khoảng từ 0.70 đến 0.75, con số chưa thực sự tốt và hiệu quả. Trong đó thì hai mô hình Logistic Regression và SVM có độ ổn định và hiệu quả cao hơn so với Multinomial Naïve Bayes khoảng 4-5% trên độ đo Accuracy. Hiệu suất của các mô hình này phụ thuộc khá nhiều vào các từ vựng trong tập huấn luyện, bởi các mô hình mà nhóm thực nghiệm chỉ sử dụng các đặc trưng và tính chất liên quan đến số lần xuất hiện của các từ, đây vừa là điểm mạnh vừa là điểm hạn chế. Điểm mạnh là ý tưởng khá đơn giản, dễ triển khai và không tốn quá nhiều tài nguyên, tuy nhiên chỉ xét đến số lần xuất hiện mà không nắm bắt được ngữ cảnh thật sự mà từ vựng đó đang xuất hiện hay thứ tự xuất hiện sẽ ảnh hưởng lớn đến hiệu suất tổng thể của mô hình, ngoài ra nếu các từ vựng trên mẫu kiểm tra không nằm trong danh sách từ vựng của quá trình huấn luyện nhưng có ảnh hưởng lớn đến cảm xúc thì sẽ dễ dàng bị loại bỏ khiến cho các mô hình có thể sai lệch.

Ngoài ra, để các mô hình có thể hoạt động tốt thì việc phân tích, thống kê và tiền xử lý dữ liệu cũng là một bước quan trọng ảnh hưởng đến hiệu quả của mô hình. Việc tiền xử lý dữ liệu sẽ loại bỏ được các trường hợp thông tin dư thừa, không cần thiết như các đường link, các mã,... không mang lại thông tin mà còn khiến cho mô hình khó học được những đặc trưng. Trong đồ án này thì nhóm đánh giá việc tiền xử lý của nhóm đạt được khoảng 80% hiệu quả, bởi có những trường hợp liên quan đến vấn đề ngữ pháp, ngữ nghĩa mà nhóm chưa xử lý được, bởi đây là văn phong đánh giá của khách hàng, họ sẽ có xu hướng viết một cách tự do, thoải mái để biểu hiện cảm xúc hay kể lại các trải nghiệm của mình, và bởi số lượng khá nhiều nên nhóm chưa tìm được giải pháp để xử lý triệt để vấn đề này. Bên cạnh đó, việc chuyển hóa các đánh giá trên thang điểm về sao sang các nhãn liên quan đến cảm xúc cũng phần nào gây ra các vấn đề tiềm ẩn, các đánh giá 1 và 5 sao thì sẽ thường hiếm có vấn đề xảy ra, nhưng các đánh giá từ 2 đến 4 sao sẽ mang lại nhiều vấn đề, bởi có những đánh giá thực sự tốt hoặc thực sự tệ nhưng

họ đánh giá nhầm mức sao tương ứng, hoặc có những trường hợp mà đánh giá 3 sao là trung tính nhưng có người lại thiên về khen ngợi, có người lại thiên về chê bai, điều này cũng gây ra khó khăn cho mô hình trong quá trình học các từ và các. Nhóm tin rằng nếu dữ liệu được kiểm định kỹ hơn, xử lý triệt để và sạch sẽ hơn thì hiệu suất của các mô hình còn có thể tăng lên đáng kể.

Sau khi hoàn thành đồ án, nhóm có cho mình một kiến thức, trải nghiệm để tiền xử lý, làm sạch dữ liệu và xây dựng các mô hình giải quyết bài toán về nhận diện cảm xúc, một bài toán kinh điển trong lĩnh vực học máy nói chung và xử lý ngôn ngữ tự nhiên nói riêng. Từ nền tảng của đồ án, nhóm dự định sẽ thực hiện và mở rộng đề tài hơn, có thể áp dụng các kỹ thuật trích xuất tự nhiên nâng cao như Word Embedding, Glove,.. hay các mô hình học sâu, ngôn ngữ lớn tiên tiến, hiện đại như LSTM, Bert, LLM để khai thác ngữ cảnh và thứ tự từ tốt hơn

TÀI LIỆU THAM KHẢO

1. An introduction to TF-IDF: What it is and How to use it -
<https://www.semrush.com/blog/tf-idf/?msockid=153655a6a79e61e1134540e1a677606c>
2. Text Classification using Logistic Regression -
<https://www.geeksforgeeks.org/machine-learning/text-classification-using-logistic-regression/>
3. Advantages and Disadvantages of Logistic Regression -
<https://www.geeksforgeeks.org/advantages-and-disadvantages-of-logistic-regression/>
4. Multinomial Naïve Bayes - <https://www.geeksforgeeks.org/machine-learning/multinomial-naive-bayes/>
5. Multinomial Naive Bayes Explained: Function, Advantages & Disadvantages, Applications - <https://www.upgrad.com/blog/multinomial-naive-bayes-explained/>
6. Implementing a Multinomial Naive Bayes Classifier from Scratch with Python -
<https://medium.com/@johnm.kovachi/implementing-a-multinomial-naive-bayes-classifier-from-scratch-with-python-e70de6a3b92e>
7. Support Vector Machine (SVM) Algorithm -
<https://www.geeksforgeeks.org/machine-learning/support-vector-machine-algorithm/>
8. What Is Sentiment Analysis? - <https://builtin.com/machine-learning/sentiment-analysis>
9. Linear SVM Classifier: Step-by-step Theoretical Explanation with Python Implementation - <https://medium.com/@tarlanahad/linear-svm-classifier-step-by-step-theoretical-explanation-with-python-implementation-d86c4973dc33>
10. Multinomial Logistic Regression With Python -
<https://machinelearningmastery.com/multinomial-logistic-regression-with-python>
11. What is text classification? - <https://aws.amazon.com/what-is/text-classification/>