

MÔ HÌNH CHIA SẺ TIN TỨC THẬT VÀ GIẢ TRÊN MẠNG XÃ HỘI

MẠNG LƯỚI

qua

Abishai Niêm vui



Một luận án

nộp một phần để hoàn thành

của các yêu cầu về mức độ

Thạc sĩ Khoa học Máy tính

Đại học Boise State

Tháng 8 năm 2021

© 2021

Abishai Niêm vui

MỌI QUYỀN ĐƯỢC BẢO LƯU

TRƯỜNG CAO ĐẲNG ĐẠI HỌC BANG BOISE

ỦY BAN QUỐC PHÒNG VÀ PHÊ DUYỆT ĐỌC CUỐI CÙNG

của luận án được nộp bởi

Abishai Niềm vui

Tên luận văn: MÔ HÌNH CHIA SẺ TIN TỨC THẬT VÀ GIẢ TRÊN MẠNG XÃ HỘI
MẠNG LƯỚI

Ngày thi vấn đáp cuối kỳ: 09 tháng 6 năm 2021

Những cá nhân sau đây đã đọc và thảo luận về luận án do sinh viên Abishai Joy nộp, và họ đã đánh giá bài thuyết trình và phản hồi của sinh viên đối với các câu hỏi trong kỳ thi vấn đáp cuối cùng. Họ thấy rằng sinh viên đã vượt qua kỳ thi vấn đáp cuối cùng.

Francesca Spezzano Tiến sĩ

Chủ tịch, Ủy ban giám sát

Jerry Alan trượt bằng tiến sĩ

Thành viên, Ủy ban giám sát

Tiến sĩ Edoardo Serra

Thành viên, Ủy ban giám sát

Sự chấp thuận đọc cuối cùng của luận án đã được Francesca Spezzano, Tiến sĩ, Chủ tịch Ủy ban giám sát, chấp thuận. Luận án đã được chấp thuận bởi Cao đẳng sau đại học.

LỜI CẢM ƠN

Tôi vô cùng biết ơn người hướng dẫn của tôi, Tiến sĩ Francesca Spezzano, nghiên cứu này công việc sẽ không thể thực hiện được nếu không có sự hướng dẫn và lời khuyên vô giá của cô ấy.

lời cảm ơn chân thành đến Tiến sĩ Edoardo Serra và Tiến sĩ Jerry Alan Fails vì sự hỗ trợ kỹ thuật của họ hỗ trợ và phản hồi về nghiên cứu của tôi. Tôi cũng muốn cảm ơn các đồng đội của tôi (Anu, Nikesh và Steven) vì sự giúp đỡ tận tình của họ. Cuối cùng, tôi muốn bày tỏ lòng biết ơn của mình với chồng tôi, cha mẹ tôi, chị gái tôi và bố mẹ chồng tôi. Không có sự giúp đỡ to lớn của họ sự hiểu biết và khuyến khích trong vài năm qua, sẽ không thể tôi hoàn thành việc học của mình. Cảm ơn Đáng toàn năng đã ban cho tôi sức khỏe và tâm trí trong suốt chuyến phiêu lưu này.

TÓM TẮT

Phương tiện truyền thông trực tuyến đang thay đổi ngành công nghiệp tin tức truyền thống và làm giảm vai trò của các nhà báo, báo chí và thậm chí các kênh tin tức. Điều này đến lượt nó lại nâng cao khả năng của tin tức giả mạo để tác động đến dư luận về các chủ đề quan trọng. Mối đe dọa của tin tức giả mạo khá là cấp bách, vì nó cho phép những người dùng có ý đồ xấu chia sẻ chương trình nghị sự của họ với một nhóm lớn hơn khán giả. Các nền tảng truyền thông xã hội lớn như Twitter, Facebook, v.v. đang tạo ra rất dễ để phát tán tin giả do việc kiểm duyệt/kiểm tra thực tế đối với những tin này rất ít nền tảng.

Công trình này nhằm mục đích dự đoán việc chia sẻ tin tức giả và thật trên mạng xã hội. Cụ thể Về cơ bản, chúng tôi sử dụng ảnh hưởng đa cấp, được rút ra từ Sự lan truyền của Đổi mới (DOI) Lý thuyết về một tập dữ liệu thực tế và dự đoán liệu một người dùng nhất định có chia sẻ hay không và khi nào thông tin trên phương tiện truyền thông xã hội. Chúng tôi đưa ra giả thuyết rằng việc chia sẻ tin tức giả và thật sẽ tốt hơn được dự đoán bằng cách xem xét các thuộc tính tính năng cấp độ người dùng, tin tức và mạng cùng nhau.

Chúng tôi cũng đang dự đoán thời gian trôi qua giữa lượt chia sẻ của người có ảnh hưởng và người theo dõi thông qua phân tích sinh tồn. Các bộ phân loại nhị phân như Support Vector Machine (SVM), Random Forest, v.v. được sử dụng để dự đoán việc chia sẻ tin tức giả và thật. Điều này cách tiếp cận được chứng minh bằng cách sử dụng một tập dữ liệu bao gồm 1.572 người dùng được lấy mẫu từ kho lưu trữ FakeNewsNet. Kết quả của chúng tôi cho thấy sự gia tăng 30% trong Khu vực Dưới Đặc điểm hoạt động của máy thu (AUROC) so với đường cơ sở tốt nhất. Việc chia sẻ tin tức thật và giả cho thấy sự phụ thuộc cao vào mức độ tương đồng của người dùng, sức mạnh liên kết,

và các tính năng rõ ràng.

Hơn nữa, phân tích cho thấy người dùng có các đặc điểm đặc trưng như yêu thích, sự tự vượt qua, lý tưởng, sự bảo tồn và sự cởi mở với sự thay đổi có xu hướng chia sẻ thực tế tin tức, trong khi người dùng có các đặc điểm nổi trội như tự nâng cao, tò mò, gần gũi, cấu trúc và sự hài hòa có nhiều khả năng chia sẻ tin giả.

Cuối cùng, phân tích sự sống còn được sử dụng để dự đoán thời gian trôi qua giữa các lần bị ảnh hưởng. encer và người theo dõi chia sẻ. Chỉ số Concordance (C-Index) cho việc chia sẻ tin tức thực sự là thấp hơn một chút so với mức cơ sở và Chỉ số C của Rừng sống sót ngẫu nhiên (RSF) có thể so sánh với đường cơ sở để chia sẻ tin tức giả mạo. Hơn nữa, khi so sánh so với các mô hình cơ sở hồi quy, Sai số tuyệt đối trung bình (MAE) là đáng kể ít hơn trong RSF khi chia sẻ cả tin tức thật và tin tức giả.

MỤC LỤC

LỜI CẢM ƠN iv

TÓM TẮT v

DANH SÁCH HÌNH ẢNH ix

DANH SÁCH BẢNG xi

DANH SÁCH CÁC CHỮ VIẾT TẮT xii

1 GIỚI THIỆU 1

2 CÔNG TRÌNH LIÊN QUAN 6

3 BỘ DỮ LIỆU. 12

 3.0.1 Mô tả tập dữ liệu. 12

 3.0.2 Tạo tập dữ liệu. 13

4 PHƯƠNG PHÁP VÀ TÍNH NĂNG 17

 4.1 Phương pháp 31

 4.1.1 Phân tích sự sống còn. 31

 4.1.2 Kiểm duyệt 32

5 THÍ NGHIỆM VÀ KẾT QUẢ 36

5.1 Bối cảnh thử nghiệm 36

5.1.1 So sánh với đường cơ sở 37

5.1.2 Phân tích việc chia sẻ tin tức thật và tin giả. 38

5.1.3 Bối cảnh thử nghiệm cho phân tích sinh tồn. 43

6 KẾT LUẬN VÀ CÔNG VIỆC TƯƠNG LAI 51

6.0.1 Công việc trong tương lai 52

6.0.2 Hạn chế 53

DANH SÁCH CÁC HÌNH ẢNH

1.1 Sơ đồ cho công việc đề xuất. 4

3.1 Cái nhìn thoáng qua về tin giả (màu đỏ) và tin thật (màu xanh) trong tập dữ liệu. 14

3.2 Phân phối tin tức thật và tin giả. 15

4.1 Phân phối tin tức thật và tin giả chia sẻ theo liên kết chính trị
của người dùng. 20

4.2 Phân bố chia sẻ tin tức giả và thật theo giới tính và độ tuổi của
người sử dụng. 21

4.3 Từ và cụm từ liên quan đến căng thẳng thấp (màu xanh) và căng thẳng cao
(màu đỏ) người dùng. 23

4.4 Kiểm duyệt đúng. 33

4.5 Cây sống sót ngẫu nhiên. 34

5.1 Đóng góp từ mỗi Nhóm tính năng. 40

5.2 Kết quả cắt bỏ tính năng (Tin tức) 41

5.3 Kết quả loại bỏ tính năng (Mạng). 42

5.4 Kết quả loại bỏ tính năng (Người dùng) 46

5.5 Kết quả loại bỏ tính năng (Người dùng) 47

5.6 Tính quan trọng của các đặc điểm tính cách đối với tin tức giả (màu đỏ) và tin tức thật
chia sẻ (màu xanh). 48

5.7 Tính năng quan trọng của các tính năng chính để chia sẻ tin tức giả (màu đỏ) và tin tức thật
(màu xanh da trời). 49

5.8 Kết quả cắt bỏ tính năng chính được đơn giản hóa. 50

DANH SÁCH CÁC BẢNG

3.1 FakeNewsNet: Chi tiết về tập dữ liệu từ phương tiện truyền thông Politifact.	13
3.2 Bộ dữ liệu tính toán - Chia sẻ tin tức giả mạo.	16
3.3 Bộ dữ liệu tính toán - Chia sẻ tin tức thực tế.	16
5.1 Kết quả phân loại: Kết quả phân loại cho hành vi Chia sẻ tin tức giả mạo. . .	39
5.2 Kết quả phân loại: Kết quả phân loại cho Real News Sharing. . .	39
5.3 Các tính năng chính: Kết quả phân loại chia sẻ tin tức giả mạo . . .	45
5.4 Các tính năng chính: Kết quả phân loại cho việc chia sẻ tin tức thực sự . . .	45
5.5 Bộ dữ liệu - Chia sẻ tin tức giả mạo	45
5.6 Bộ dữ liệu - Chia sẻ tin tức thực tế	45
5.7 Kết quả hồi quy: Kết quả hồi quy cho việc chia sẻ tin tức thực tế. . .	45
5.8 Kết quả hồi quy: Kết quả hồi quy cho việc chia sẻ tin tức giả mạo. . .	45

DANH SÁCH CÁC TỪ VIẾT TẮT

Khu vực AUROC dưới Đặc điểm hoạt động của máy thu

Chỉ số C-Index Concordance

DOI Lý thuyết khuếch tán đổi mới

Mô hình thác đổ độc lập ICM

LIWC Ngôn ngữ học và số lượng từ

Mô hình ngưỡng tuyến tính LTM

MAE Sai số tuyệt đối trung bình

Rừng sinh tồn ngẫu nhiên RSF

Kỹ thuật lấy mẫu quá mức thiểu số tổng hợp SMOTE

Máy hỗ trợ vectơ SVM

Tỷ lệ người theo dõi-người theo dõi trên Twitter của TFF

CHƯƠNG 1:

GIỚI THIỆU

Học sinh trung học ở Philadelphia tin rằng trái đất phẳng vì ý tưởng

họ đã học được điều này từ ngôi sao bóng rổ Kyrie Irving, người đã nói điều đó trên một podcast [1].

Câu chuyện có vẻ buồn cười này làm nổi bật tác động đáng báo động của thông tin sai lệch trong

thời đại kỹ thuật số ngày nay. Từ điển Cambridge định nghĩa tin giả là những câu chuyện sai sự thật

có vẻ như là tin tức, lan truyền trên internet hoặc sử dụng các phương tiện truyền thông khác, thường được tạo ra để

ảnh hưởng đến quan điểm chính trị hoặc như một trò đùa [2]. Tin tức giả thường được đặc trưng bởi việc sử dụng

có nội dung, nguồn hoặc xuất xứ không thể xác minh được, thường được thiết kế để thu hút cảm xúc. Web

Công nghệ 2.0 đã đẩy nhanh hơn nữa việc tiêu thụ tin tức giả mạo và tác động của nó đối với

xã hội, dẫn đến nhu cầu cấp thiết phải hiểu được sự lan truyền của thông tin sai lệch trong

phương tiện truyền thông xã hội. Một trong những nền tảng truyền thông xã hội phổ biến nhất hiện nay là Twitter. Như

theo số liệu thống kê của Twitter, khoảng 145 triệu người dùng hoạt động hàng ngày (trong số 330 triệu)

dành trung bình 3,39 phút mỗi phiên trên các mạng xã hội trực tuyến và

ước tính chia sẻ 500 triệu tweet mỗi ngày [3]. Trong khi một phần nhỏ

người dùng tích cực truyền bá thông tin giả mạo, các nền tảng lớn như Twitter khuếch đại

tác động dựa trên động lực chia sẻ. Hiệu động lực của tin tức

chia sẻ (đặc biệt là tin giả) sẽ cung cấp một phương tiện để ngăn chặn việc phát tán

tin giả trước hoặc ít nhất là nhanh chóng khi phát hiện.

Phương tiện truyền thông in ấn truyền thống liên quan đến nghiên cứu, điều tra, đánh giá ngang hàng và

không thể dễ dàng chỉnh sửa sau khi xuất bản. Tuy nhiên, tin tức trên phương tiện truyền thông xã hội thường không được kiểm duyệt, có thể bị thay đổi bởi bất kỳ ai và tốc độ phát tán có thể thay đổi động tại bất kỳ thời điểm nào. Hiểu được mối tương quan giữa nội dung tin tức, hành vi (hoặc đặc điểm) của người dùng và tốc độ lan truyền thông tin sẽ giúp ngăn ngừa lan truyền tin giả. Các tính năng tin tức như phong cách, sự phức tạp và tâm lý; người dùng các đặc điểm như tính cách cá nhân, khuynh hướng chính trị, nhân khẩu học và thời gian giữa việc chia sẻ tin tức có thể cung cấp những hiểu biết hữu ích để mô hình hóa tin tức trực tuyến (và giả mạo tin tức) lan truyền. Thời gian của một tweet có thể đặc biệt hữu ích trong việc nghiên cứu chia sẻ tin tức giả và phát hiện tin tức giả. Như chúng tôi đã chỉ ra trước đó, tin tức giả cố gắng khơi gợi cảm xúc và điều này thường dẫn đến việc truyền bá tin tức nhanh hơn. Nghiên cứu này khám phá những mối tương quan này để mô hình hóa việc chia sẻ tin tức thật và tin tức giả.

Các mô hình cổ điển cho sự khuếch tán thông tin như Độc lập Cascade và Tuyến tính Các mô hình ngưỡng giả định rằng người dùng sẽ chia sẻ tin tức chỉ với một số xác suất theo thực tế là một số người bạn của họ trước đó đã chia sẻ tin tức tương tự [4]. Tuy nhiên, các công trình gần đây về việc chia sẻ tin tức giả trong lĩnh vực khoa học xã hội đã chỉ ra rằng quyết định chia sẻ hay không chia sẻ một tin tức nào đó của người dùng không chỉ phụ thuộc vào ảnh hưởng của bạn bè mà còn phụ thuộc vào đặc điểm cụ thể của người dùng (ví dụ: thông tin nhân khẩu học, thuộc tính hồ sơ Twitter, hành vi và hoạt động trên Twitter, v.v.), tin tức nhận được (ví dụ: tiêu đề và nội dung, v.v.) và bối cảnh xã hội (ví dụ: số lượng người theo dõi và theo dõi, sức mạnh liên kết, v.v.) [5]. Tất cả các khía cạnh này đều phù hợp với những gì được lý thuyết hóa bởi lý thuyết khuếch tán đổi mới để giải thích cách một đổi mới (trong trường hợp của chúng tôi là tin tức) lan truyền trong mạng xã hội [6]. Hơn nữa, trong cuộc sống thực, người dùng không chia sẻ tin tức ngay khi nhận được, nhưng có một số người dùng tiếp xúc với tin tức trong một thời gian trước khi chia sẻ, trong khi những người khác trở nên hoài nghi và làm

không chia sẻ gì cả.

Vì vậy, mục tiêu chính của chúng tôi thông qua luận án này là dự đoán liệu và khi nào người dùng sẽ chia sẻ một tin tức từ người có sức ảnh hưởng của mình. Đối với điều này, chúng tôi đang áp dụng sự lan truyền của lý thuyết đổi mới để mô hình hóa cách tin tức thật và tin giả được chia sẻ trên Twitter người dùng. Với bằng chứng từ các nghiên cứu khoa học xã hội, chúng tôi đưa ra giả thuyết rằng thật và giả chia sẻ tin tức được dự đoán tốt hơn khi người dùng, tin tức và đặc điểm mạng xã hội tất cả đều được tính đến [5]. Tất cả các yếu tố này chưa bao giờ được kết hợp thành một mô hình dự đoán độc đáo hoặc đã được thử nghiệm trên quy mô lớn trước đây.

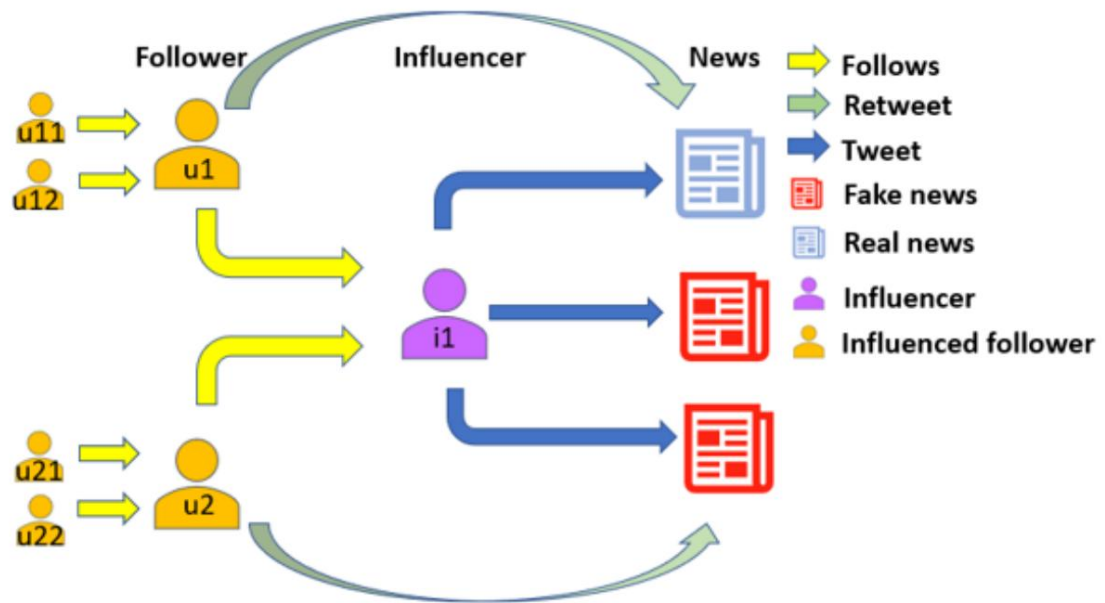
Cụ thể, chúng tôi sẽ giải quyết các vấn đề sau:

- cho rằng người dùng 'u1' bị ảnh hưởng bởi một số tin tức nhất định 'n' bởi ít nhất một trong những người có sức ảnh hưởng đến họ 'i1' (tức là, 'u1' đang theo dõi 'i1' và 'i1' đã chia sẻ một số tin tức 'n' trong số những người theo dõi họ), dự đoán liệu người dùng 'u1' có chia sẻ tin tức 'n' trong số những người theo dõi họ (u11 & u12); và
- cho rằng người dùng 'u1' bị ảnh hưởng bởi một số tin tức nhất định 'n' bởi ít nhất một trong những người có sức ảnh hưởng 'i1' của họ, dự đoán thời gian trôi qua giữa thời điểm 'i1' chia sẻ tin tức 'n' và người dùng 'u1' cũng sẽ chia sẻ 'n' cho những người theo dõi họ (u11 và u12).

Như chúng tôi đã đề cập trước đó, Twitter là một trong những phương tiện truyền thông xã hội phổ biến nhất nền tảng. Do đó, chúng tôi sử dụng bộ dữ liệu Twitter từ FakeNewsNet1 cho phân tích của mình. Bộ dữ liệu bao gồm nội dung tin tức, bối cảnh xã hội và thông tin không gian-thời gian. Nó chứa dữ liệu như dữ liệu hồ sơ Twitter, dòng thời gian tweet có nội dung, tiêu đề tin tức như cũng như thông tin nội dung và mối quan hệ giữa người theo dõi.

Mục tiêu chung là dự đoán liệu người dùng có chia sẻ nội dung hay không và khi nào vòng tròn xã hội của anh ấy/cô ấy. Hình 1.1 phác thảo phương pháp luận mà chúng tôi triển khai

1<https://github.com/KaiDMML/FakeNewsNet/tree/master/code>



Hình 1.1: Sơ đồ cho công việc đề xuất

để dự đoán hành vi chia sẻ tin tức của người dùng. Dữ liệu thô bao gồm thông tin từ Hồ sơ Twitter, tweet, nội dung tin tức và thông tin chi tiết về người theo dõi. Ảnh hưởng đa cấp cách tiếp cận, được rút ra từ lý thuyết khuếch tán đổi mới (DOI), được sử dụng để giải thích làm thế nào, tại sao và ở tốc độ nào những ý tưởng và công nghệ mới lan truyền qua một dân số hoặc hệ thống xã hội [6]. Dữ liệu được chia thành ba loại (người dùng, tin tức, và mạng) để thực hiện phương pháp tiếp cận ảnh hưởng đa cấp. Dữ liệu được tiếp tục được chuyển đổi và sử dụng để tính toán các tính năng như hành vi, ngôn ngữ và mối quan hệ của người dùng sức mạnh. Những đặc điểm này được đưa thêm vào các bộ phân loại và mô hình sống sót. Đối với phần đầu tiên trong nghiên cứu của chúng tôi, các bộ phân loại như hồi quy logistic, cây bổ sung, rừng ngẫu nhiên và máy vectơ hỗ trợ được sử dụng và đánh giá bằng cách sử dụng Area Under Đặc điểm hoạt động của máy thu (AUROC) và độ chính xác trung bình. Thứ hai một phần tập trung vào việc dự đoán thời gian người dùng chia sẻ tin tức (thời gian chia sẻ).

Theo Vosoughi và cộng sự, tin tức giả lan truyền nhanh hơn, xa hơn và rộng hơn tin tức thực sự [7]. Do đó, thời gian chia sẻ là một đặc điểm quan trọng của tin tức giả mạo và biết tốc độ lan truyền là quan trọng để hiểu thời gian phản ứng cần thiết để ngăn chặn sự lan truyền tin tức giả mạo. Do đó, thời gian là thông tin quan trọng có thể giúp ích với cả việc phát hiện và ngăn chặn việc chia sẻ tin tức giả mạo. Phân tích này đòi hỏi tính toán thời gian trôi qua giữa các tweet và retweet. Hơn nữa, ngẫu nhiên Rừng sinh tồn được sử dụng để dự đoán thời gian xảy ra sự kiện.

Kết quả của chúng tôi cho thấy việc chia sẻ tin tức được dự đoán tốt hơn khi có nhiều tính năng được xem xét. Trong số các tính năng, các tính năng dựa trên tin tức đã vượt trội hơn, tiếp theo là thuộc tính người dùng và mạng. Phương pháp kết hợp tin tức, mạng và người dùng Các tính năng đã tăng tổng thể AUROC lên 30% so với mức cơ sở tốt nhất. Chúng tôi sử dụng mô hình khuếch tán Ngưỡng tuyến tính và Tầng độc lập làm cơ sở. Các mô hình này sử dụng xác suất lan truyền để lây nhiễm cho người dùng hoặc người theo dõi không hoạt động và không phụ thuộc vào tin tức hoặc tính năng của người dùng. Mô hình sinh tồn được đánh giá bằng cách sử dụng chỉ số phù hợp và sai số tuyệt đối trung bình. Các phát hiện cho thấy rằng sai số tuyệt đối thu được từ mô hình sống sót ít hơn đáng kể và đáng tin cậy hơn để chia sẻ tin tức giả mạo.

Luận văn được tổ chức như sau, Chương 2 báo cáo công trình liên quan. Chương 3 mô tả tập dữ liệu được sử dụng cho các thí nghiệm và phân tích. Chương 4 giải thích các tính năng được tính toán và các phương pháp. Kết quả phân tích của chúng tôi được lập bảng trong Chương 5, Thí nghiệm và Kết quả. Cuối cùng, Chương 6 đưa ra kết luận cùng với công việc trong tương lai.

CHƯƠNG 2:

CÔNG VIỆC LIÊN QUAN

Các nghiên cứu trước đây cho thấy người dùng điển hình chia sẻ tin tức có thể được coi là người dẫn đầu dư luận hoặc người có ảnh hưởng. Người dẫn đầu dư luận áp dụng những ý tưởng mới và thường được mạng lưới của họ tiếp cận để xin lời khuyên hoặc thông tin về nội dung được chia sẻ bởi người có sức ảnh hưởng. Mặc dù không có lý thuyết toàn cầu nào để nghiên cứu việc chia sẻ tin tức trên mạng xã hội các tác phẩm, nghiên cứu cho thấy nhiều tác phẩm trước đây rút ra suy luận từ lý thuyết DOI [5]. Bên cạnh DOI, các lý thuyết về ảnh hưởng xã hội và các khái niệm về tính tương tác, sự tham gia thực tế và cách tiếp cận sử dụng và thỏa mãn có phần liên quan [5]. Trong tất cả các lý thuyết này, việc tìm kiếm địa vị, đạt được danh tiếng và thu hút sự chú ý của mọi người sự chú ý đến quan điểm và ý tưởng của riêng mình là động lực chính của việc chia sẻ tin tức [8]. Shu et al. [9, 10] tập trung vào việc xem xét mối quan hệ ba bên giữa người dùng, nhà xuất bản và nội dung tin tức. Sự tham gia của người dùng thể hiện quá trình lan truyền tin tức theo thời gian, cung cấp thông tin hỗ trợ hữu ích để suy ra tính xác thực của các bài báo. Shu et al. [11] đã đề xuất một phương pháp phát hiện sử dụng các tính năng rõ ràng và ngầm định từ dữ liệu, có khả năng phân biệt tin tức giả mạo. Một số ví dụ bao gồm đăng ký thời gian và thông tin nhân khẩu học cho các tính năng rõ ràng và ẩn tương ứng. Đăng ký là thời điểm người dùng đăng ký trên nền tảng mạng xã hội, điều này thông tin được bao gồm trong dữ liệu hồ sơ người dùng từ Twitter. Các thông tin nhân khẩu học là được tính toán bằng cách sử dụng id, tên, tên màn hình, mô tả, ngôn ngữ, đường dẫn hình ảnh và

thay đổi kích thước ảnh đại diện của người dùng.

Các nghiên cứu gần đây cũng suy ra rằng mọi người có xu hướng chia sẻ thông tin sai lệch trước đây được gắn thẻ là không chính xác. Cuộc khảo sát do Pennycook và cộng sự [12] thực hiện cho thấy rằng tiêu đề nhận được đánh giá độ chính xác cao hơn của con người so với tiêu đề sai. Tuy nhiên, người dùng trong nghiên cứu đó dường như hầu như không tính đến thông tin này khi cân nhắc những gì để chia sẻ trên phương tiện truyền thông xã hội. Những lý do chính đằng sau hành vi này là (1) thu hút nhiều người theo dõi hơn, (2) báo hiệu tư cách thành viên nhóm của một người và (3) thu hút với nội dung gợi cảm xúc khiến khán giả mất tập trung vào tính xác thực của tin tức.

Như đã đề cập trong Chương 1, việc xem xét nhiều cấp độ ảnh hưởng là một giải pháp hiệu quả cách nghiên cứu hành vi chia sẻ tin tức trên phương tiện truyền thông xã hội [13]. Nhiều cấp độ của ảnh hưởng bao gồm mạng lưới khuếch tán, ảnh hưởng cá nhân và các thuộc tính đối mới (thuộc tính tin tức). Ma et al. [13] cũng giải thích tầm quan trọng của việc phân tích mạnh mẽ và mối quan hệ yếu cùng với số lượng người theo dõi trong một vòng tròn xã hội. Không phải tất cả các nhân khẩu học các thuộc tính là những yếu tố dự báo cho việc chia sẻ thông tin trên phương tiện truyền thông xã hội. Nó đã được chứng minh rằng việc thiếu hiểu biết về phương tiện truyền thông kỹ thuật số là một trong những lý do khiến tin tức giả mạo gia tăng chia sẻ [14]. Điều này cho thấy rằng độ tuổi của người dùng là một yếu tố dự báo tốt để hiểu chia sẻ tin tức giả trên mạng xã hội. Hơn nữa, không có biến số nhân khẩu học nào khác - giới tính, chủng tộc, trình độ học vấn và thu nhập - có tác động dự đoán mạnh mẽ đến việc chia sẻ thông tin giả tin tức [14]. Nghiên cứu của Yaqub và cộng sự cho thấy việc bổ sung chỉ số độ tin cậy nội dung truyền thông xã hội có khả năng làm tăng khả năng hiểu biết thông tin [15]. Tuy nhiên, phát hiện của tác giả cho thấy hiệu quả của chúng thay đổi tùy theo loại của chỉ số và đặc điểm cá nhân của người dùng [15]. Bài báo của Yaqub et al. cũng thảo luận về các lý do khác nhau đằng sau ý định chia sẻ [15]. Dưới sự kiểm soát

điều kiện (khi người dùng không được cung cấp chỉ số độ tin cậy), động cơ chính để người dùng chia sẻ hoặc không chia sẻ tin tức đúng hay sai là: (1) người dùng muốn / không muốn để đưa tin tức thú vị/không thú vị vào trang mạng xã hội của họ, (2) tin tức sẽ / sẽ không kích hoạt một cuộc thảo luận giữa những người bạn, (3) người dùng muốn/không muốn chia sẻ một tin tức đúng/sai, và (4) người dùng muốn/không muốn chia sẻ tin tức này vì nó là/là không liên quan đến cuộc sống của anh ấy hoặc cô ấy [15]. Khi người dùng được cung cấp một chỉ số độ tin cậy, những người tham gia khảo sát đã chọn không chia sẻ tin tức giả mạo 30% thời gian. Những kết quả có thể bị thiên vị vì người dùng biết độ tin cậy của tin tức trước khi khảo sát. Hơn nữa, số liệu thống kê được thu thập thủ công và không có phân tích tự động, có thể mở rộng đã được phát triển [15].

Vosoughi et al. [7] đã khám phá động lực khuếch tán của tin tức đúng và tin tức sai. Theo nghiên cứu này, tin giả lan truyền nhanh hơn, xa hơn, sâu hơn và rộng hơn đáng kể hơn tin tức thực sự. Phân tích cho thấy tin tức về chính trị và truyền thuyết đô thị là lan truyền nhiều nhất. Bài viết này cũng nhấn mạnh rằng việc loại bỏ bot bằng cách sử dụng bot thuật toán phát hiện không bao giờ thay đổi kết quả của chúng về động lực khuếch tán.

Những nỗ lực nghiên cứu hiện tại khai thác nhiều tính năng khác nhau của dữ liệu, bao gồm mạng lưới tính năng. Bài báo của Shu et al. [16] cũng giải thích rằng nguồn gốc của tin tức giả chỉ ra những người khởi tạo. Nguồn gốc có thể giúp trả lời các câu hỏi như liệu tin tức đã được sửa đổi trong quá trình truyền bá và người tạo ra nó thông tin được kết nối với việc truyền tải tuyên bố. Tuy nhiên, nó không giúp dự đoán liệu người dùng có chia sẻ hay tốc độ chia sẻ tin tức giả mạo hay không.

Phân tích mạng xã hội đang ngày càng phổ biến và là một trong những nghiên cứu quan trọng khu vực tìm kiếm trong lĩnh vực này là Phổ biến thông tin. Có hai cách được sử dụng rộng rãi các mô hình khuếch tán thông tin, cụ thể là (i) Mô hình ngưỡng khuếch tán và (ii) Cas-

Mô hình khuếch tán.

Mô hình thác nước độc lập (ICM) là một mô hình khuếch tán thông tin ngẫu nhiên nơi thông tin chảy qua mạng thông qua một thác nước. Các nút có thể có hai trạng thái, (i) Hoạt động: Có nghĩa là nút đã bị ảnh hưởng bởi thông tin trong sự khuếch tán. (ii) Không hoạt động: Nút không biết thông tin hoặc không bị ảnh hưởng bởi thông tin trong quá trình khuếch tán.

Quá trình này diễn ra theo các bước riêng biệt. Vào lúc bắt đầu quá trình ICM, một vài các nút, được gọi là các nút hạt giống, đã chia sẻ tin tức. Khi nhận được thông tin, các nút này trở nên hoạt động. Trong mỗi bước rời rạc, một nút hoạt động cố gắng ảnh hưởng đến một trong những người hàng xóm không hoạt động của nó. Cùng một nút sẽ không bao giờ có cơ hội khác để kích hoạt cùng một người hàng xóm không hoạt động. Sự thành công phụ thuộc vào sự lan truyền xác suất hòa của họ. Xác suất lan truyền của một trận hòa là xác suất của cái nào có thể ảnh hưởng đến nút kia. Trong thực tế, xác suất lan truyền là mối quan hệ phụ thuộc, nghĩa là mỗi cạnh sẽ có một giá trị khác nhau. Quá trình kết thúc khi không có nút nào được kích hoạt từ trạng thái không hoạt động [4].

Trong Mô hình ngưỡng tuyến tính (LTM), một nút v bị ảnh hưởng bởi mỗi hàng xóm w theo trọng số $b_{v,w}$. Động lực của quá trình sau đó diễn ra như sau. Mỗi nút v chọn một ngưỡng θ_v một cách ngẫu nhiên từ khoảng $[0, 1]$; điều này biểu thị phần có trọng số của các hàng xóm của v phải trở nên hoạt động để v để trở nên hoạt động. Với sự lựa chọn ngẫu nhiên các ngưỡng và một tập hợp ban đầu các hoạt động các nút A (với tất cả các nút khác không hoạt động), quá trình khuếch tán diễn ra một cách xác định trong các bước rời rạc: trong bước t , tất cả các nút hoạt động trong bước $t-1$ vẫn hoạt động và chúng tôi kích hoạt bất kỳ nút v nào mà tổng trọng số của các nút lân cận đang hoạt động của nó ít nhất là V :

$$b_{v,w} \leq 1$$

tổng của hàng xóm v

Vì vậy, ngưỡng Θ_v trực quan biểu diễn các xu hướng tiềm ẩn khác nhau của

các nút áp dụng đổi mới khi các nút lân cận của chúng thực hiện [4].

Trong một mô hình dịch tễ học như SEIR [17] (Có thể mắc bệnh-Phơi nhiễm-Nhiễm bệnh-

Đã phục hồi), một khoảng thời gian tiềm ẩn được đưa vào để cho phép thời gian chờ ngẫu nhiên trước khi

sự khởi đầu của sự lây nhiễm. Về mặt sinh học, điều này tương ứng với một giai đoạn trong

trong đó nhiễm trùng đang tự thiết lập trong vật chủ của nó nhưng không thể nhảy sang vật chủ khác

máy chủ. Khoảng thời gian tiếp xúc được cho là được phân phối giống hệt nhau cho tất cả các cặp ij

trong đó i lây nhiễm j , điều này không thực tế và khoảng cách tiếp xúc có thể thay đổi đối với các

người dùng [17]. Tác động của các biến thể phụ trợ (tính năng) lên quá trình lây nhiễm là một

mối quan tâm chính mà hầu hết các nghiên cứu trước đây chưa đề cập đến. Các mô hình khác

giống như mô hình thác nước độc lập và mô hình ngưỡng tuyến tính đã được đề xuất để

hiểu được hướng ảnh hưởng [18]. Tuy nhiên, các chi tiết của người dùng hoặc tin tức

nội dung vẫn chưa được phân tích bằng cách sử dụng các mô hình này. Chúng tôi đã sử dụng Independent

Các mô hình Ngưỡng tuyến tính và Cascade làm cơ sở cho nghiên cứu này và được so sánh

hiệu suất của phương pháp chúng tôi đề xuất.

Để khắc phục những hạn chế nêu trên, công trình của chúng tôi nghiên cứu sự khuếch tán

của lý thuyết đổi mới để khám phá cấp độ cá nhân, cấp độ mạng lưới và thuộc tính tin tức

mức độ tác động đến quyết định chia sẻ của người dùng. Hơn nữa, chúng tôi đã tiến hành một

phân tích sử dụng rừng sống sót ngẫu nhiên để dự đoán thời gian chia sẻ cho một ảnh hưởng nhất định

người dùng. Một khu rừng sống sót ngẫu nhiên là một ước tính siêu dữ liệu phù hợp với một số cây sống sót

trên các mẫu phụ khác nhau của tập dữ liệu và sử dụng phương pháp trung bình để cải thiện khả năng dự đoán

độ chính xác và kiểm soát quá mức. Phương pháp này được sử dụng để sống sót kiểm duyệt bên phải

dữ liệu như mô tả trong Phần 4.1.1.

CHƯƠNG 3:

TẬP DỮ LIỆU

3.0.1 Mô tả tập dữ liệu

FakeNewsNet là một kho lưu trữ dữ liệu đa chiều hiện có chứa hai

các tập dữ liệu có nội dung tin tức, bối cảnh xã hội và thông tin không gian thời gian [19].

tập dữ liệu được xây dựng bằng cách sử dụng hệ thống đầu cuối, FakeNewsTracker [19].

kho lưu trữ FakeNewsNet có cấu trúc có tiềm năng thúc đẩy việc nghiên cứu nhiều loại mở

nghiên cứu các vấn đề liên quan đến nghiên cứu tin tức giả mạo. Bởi vì việc chia sẻ dữ liệu Twitter

chính sách, nó chỉ chia sẻ các bài viết tin tức và ID tweet như một phần của tập dữ liệu này và

cung cấp mã để tải xuống thông tin chi tiết đầy đủ về tweet, tương tác xã hội và mạng xã hội

mạng lưới.

Kho lưu trữ mã có thể được sử dụng để tải xuống các bài báo từ các trang web đã xuất bản

và dữ liệu phương tiện truyền thông xã hội có liên quan từ Twitter. Các tập lệnh sử dụng các khóa từ

tập khóa tweet, được kích hoạt từ tài khoản nhà phát triển Twitter. Tóm tắt

của tập dữ liệu có thể được thấy trong Bảng 3.1.

Bộ dữ liệu được tải xuống từ FakeNewsNet chứa thông tin thực tế được thu thập

từ PolitiFact. PolitiFact là một trang web được điều hành bởi Tampa Bay Times, nơi

người khuôn vác và biên tập viên từ các phương tiện truyền thông kiểm tra thực tế các bài báo. Trang web xuất bản

tuyên bố ban đầu của bài báo và kết quả kiểm tra thực tế của chúng.

Bảng 3.1: FakeNewsNet: Chi tiết về tập dữ liệu từ phương tiện truyền thông Politifact

	Chính trị hiện vật
Số lượng người dùng	281.596
Tin Thật + Tin Giả	560 + 432
Số lượng tin tức	992
Số lượng Tweet	438,504
Số lượng Retweet	619,239

Có 560 tin tức thật và 432 tin tức giả trong FakeNewsNet. Các đám mây từ

đối với các thành phần nội dung tin tức của các tập dữ liệu được trình bày trong Hình 3.1.

các đám mây từ đại diện cho các từ thường xuyên xuất hiện trong một văn bản, trong đó kích thước của từ là

tỷ lệ thuận với số lần sử dụng từ ngữ. Chúng tôi quan sát thấy rằng tin tức giả

từ tập dữ liệu Politifact có nhiều nội dung chính trị hơn so với tin tức thực tế.

thông tin bối cảnh xã hội của 281K người dùng bao gồm các bài đăng của họ, hành vi của người dùng như

trả lời, đăng lại và thích cũng như thông tin siêu dữ liệu cho hồ sơ người dùng, người dùng

bài đăng và thông tin mạng xã hội. Bối cảnh động bao gồm thông tin

chẳng hạn như dấu thời gian tương tác của người dùng.

Tweet và retweet là phương tiện phổ biến để chia sẻ nội dung trên Twitter. Từ

FakeNewsNet, chúng tôi đã có 438.000 tweet và 619.000 tweet lại với thông tin chi tiết về người dùng và

thời điểm chia sẻ. Hình 3.2 cho thấy sự phân bố cổ phiếu cho cổ phiếu thật và giả

tin tức trong tập dữ liệu. Rõ ràng từ hình ảnh rằng tin tức thực sự được chia sẻ nhiều hơn

hơn và được tweet lại thường xuyên hơn tin tức giả mạo. Điều này không có gì đáng ngạc nhiên vì

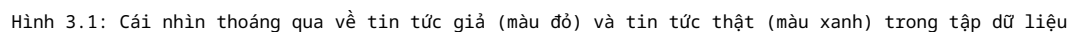
nguyên cứu của Guess và cộng sự cho thấy việc chia sẻ các bài viết từ các miền tin tức giả mạo là một

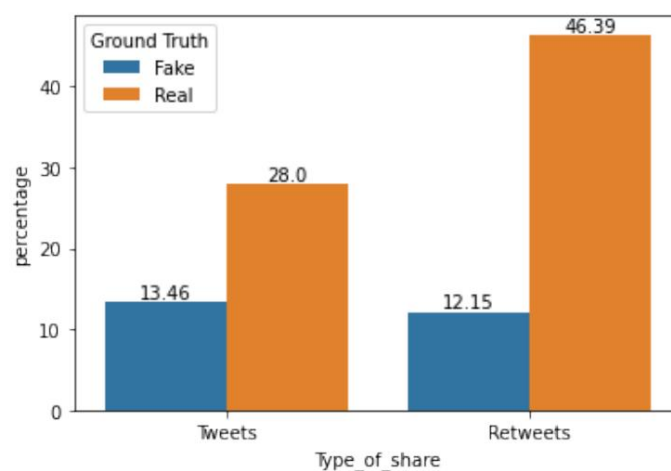
hiện tượng hiếm gặp [14], nhưng nó có thể có tác động lớn.

3.0.2 Tạo tập dữ liệu

Sau khi phân tích thăm dò, chúng tôi đã hợp nhất các tập dữ liệu từ phương tiện truyền thông Politifact vào

xác định cặp người có ảnh hưởng và người dùng bị ảnh hưởng. Cặp người dùng là cặp người dùng





Hình 3.2: Phân phối cổ phiếu tin tức thật và tin tức giả

đã tweet và retweet một tin tức nhất định. Người dùng bị ảnh hưởng là người theo dõi

người dùng có sức ảnh hưởng. Đối với mỗi người dùng có sức ảnh hưởng, ít nhất 5 trường hợp chia sẻ tin tức

được xem xét theo thứ tự thời gian. Hai tin tức được chia sẻ gần đây nhất là

được sử dụng trong việc tạo ra tập dữ liệu và các chia sẻ còn lại được nối lại với nhau

tính toán các tính năng cho nhiệm vụ phân loại, được trình bày chi tiết trong Chương 4. Trong

tập dữ liệu cuối cùng, có 2.403 người dùng bị ảnh hưởng và mỗi hàng xác định một người dùng

(người có ảnh hưởng) tiếp xúc với một tin tức và có ít nhất một người theo dõi. Nếu

Người theo dõi (người dùng bị ảnh hưởng) chia sẻ nội dung, chúng tôi gán nhãn trường hợp này là 1, nếu không thì là 0.

Việc tính toán các tính năng dựa trên người dùng đã sử dụng các dòng tweet theo dòng thời gian của người dùng bị ảnh hưởng,

được thu thập riêng biệt từ Twitter. Đối với điều này, chúng tôi đã lọc các dòng tweet trên dòng thời gian của

người dùng bị ảnh hưởng, được đăng trong khoảng thời gian của ngày xuất bản tin tức

được sử dụng trong tập dữ liệu chính cho mỗi người dùng bị ảnh hưởng. Chúng tôi chỉ có dữ liệu dòng thời gian trên 1.572

ảnh hưởng đến người dùng làm giảm tổng số người dùng. Bộ dữ liệu của chúng tôi không cân bằng

với 3.144 và 14.936 trường hợp được gán nhãn là 1 (người dùng bị ảnh hưởng) và được gán nhãn là 0 (không

người dùng bị ảnh hưởng). Chi tiết về tập dữ liệu được tính toán được đưa ra trong Bảng 3.2 và Bảng

Bảng 3.2: Bộ dữ liệu tính toán - Chia sẻ tin tức giả

	Chính trị hiện vật
Số lượng người dùng 1.557	
Số lượng tin giả 169	
Người dùng bị ảnh hưởng 527	
Không bị ảnh hưởng Người dùng 7,134	

Bảng 3.3: Bộ dữ liệu tính toán - Chia sẻ tin tức thực tế

	Chính trị hiện vật
Số lượng người dùng 1.572	
Đếm tin tức thực tế 127	
Người dùng bị ảnh hưởng 2.617	
Không bị ảnh hưởng Người dùng 7.802	

3.3 cho việc chia sẻ tin tức giả và tin tức thật, tương ứng. Cuối cùng, thời gian trôi qua giữa các tweet và retweet của cùng một tin tức cho các cặp người dùng khác nhau đã được tính toán để cung cấp sự thật cơ bản cho vấn đề dự đoán thời điểm người dùng sẽ chia sẻ thông tin như đã thảo luận trong Phần 4.1.1.

CHƯƠNG 4:

PHƯƠNG PHÁP VÀ ĐẶC ĐIỂM

Như đã thảo luận trong Chương 1, chúng tôi đã sử dụng ảnh hưởng đa cấp được rút ra từ DOI lý thuyết. Nghiên cứu này tập trung vào ba loại tính năng, cụ thể là dựa trên người dùng, tin tức-dựa trên các tính năng dựa trên mạng xã hội. Trong khi các thuộc tính rõ ràng và ngầm định trở thành một tính năng cho mỗi người dùng trong tập dữ liệu, đặc điểm về phong cách và độ phức tạp tạo thành các tính năng dựa trên tin tức. Chương này mô tả tập hợp các tính năng mà chúng tôi được sử dụng trong nghiên cứu để phân tích việc chia sẻ tin tức thật và giả.

Các tính năng liên quan đến người dùng:

- Đặc điểm tính cách

Dịch vụ IBM Watson Personality Insights sử dụng phân tích ngôn ngữ để suy ra đặc điểm tính cách nội tại của cá nhân, bao gồm tính cách Big Five đặc điểm, nhu cầu và giá trị, từ các phương tiện truyền thông kỹ thuật số như phương tiện truyền thông xã hội bài đăng. Trong nghiên cứu này, tất cả các dòng thời gian tweet được nối lại cho một ảnh hưởng đến người dùng để tính toán các đặc điểm tính cách của họ. Các tính năng com-được cung cấp bởi dịch vụ này được trình bày chi tiết như sau:

Năm lớn

Năm đặc điểm tính cách lớn, còn được gọi là mô hình năm yếu tố (FFM) và

Mô hình OCEAN là một phân loại được sử dụng rộng rãi để mô tả tính cách của con người

đặc điểm [20]. Năm chiều kích tính cách cơ bản được mô tả bởi phân loại này là sự cởi mở để trải nghiệm, sự tận tâm, sự hướng ngoại, sự dễ chịu và tính cách thần kinh. Đối với mỗi chiều kích tính cách, IBM Watson Personality Insights cũng cung cấp một bộ sáu tính năng khía cạnh bổ sung. Ví dụ, sự dễ chịu các khía cạnh bao gồm lòng vị tha, sự hợp tác, sự khiêm tốn, đạo đức, sự cảm thông và sự tin tưởng.

Nhu cầu

Các tính năng này mô tả nhu cầu của người dùng được suy ra từ văn bản họ viết và bao gồm sự phấn khích, sự hòa hợp, sự tò mò, lý tưởng, sự gần gũi, sự tự thể hiện, sự tự do, tình yêu, tính thực tế, sự ổn định, thử thách và cấu trúc.

Giá trị

Những đặc điểm này mô tả các yếu tố thúc đẩy ảnh hưởng đến quyết định của một người. làm. Chúng bao gồm sự tự vượt qua, bảo tồn, chủ nghĩa khoái lạc, tự nâng cao, và cởi mở với sự thay đổi.

• Tính năng rõ ràng

Được bảo vệ, được xác minh và thời gian đăng ký - Được bảo vệ, khi đúng, cho biết rằng điều này người dùng đã chọn bảo vệ tweet của họ. Đã xác minh cho biết liệu đó có phải là tweet đã xác minh hay không người dùng. Thời gian đăng ký là số ngày đã trôi qua kể từ ngày tạo. Những là những đặc điểm rõ ràng được Shu và cộng sự giải thích có thể mô tả người dùng tốt hơn đặc điểm [11].

Số lượng trạng thái và số lượng ủng hộ - Số lượng trạng thái cho biết số lượng tweet

(bao gồm cả các lượt retweet) do người dùng phát hành và số lượt thích cho biết số lượng

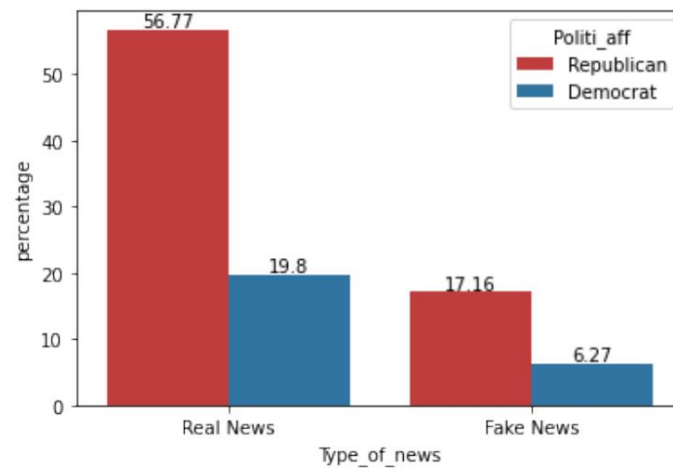
tweet mà người dùng này đã thích trong suốt thời gian tồn tại của tài khoản [11]. Các tính năng này đại diện cho mức độ hoạt động của người dùng trên mạng xã hội.

• Ý thức hệ chính trị

Điều này cho thấy sự liên kết chính trị của người dùng nhất định và liên quan đến tính toán của điểm cực từ các hashtag của người dùng bằng cách sử dụng thuật toán lựa chọn tính năng [21]. Đối với việc tính toán điểm cực, chúng tôi đã xem xét tập dữ liệu chính trị từ bài báo của Chamberlain và cộng sự, vì phạm vi thời gian của nó phù hợp với chúng tôi tập dữ liệu chính [22]. Tập dữ liệu có các tweet và các hashtag liên quan của các chính trị gia. Là một phần của quá trình, TfidfVectorizer đã được sử dụng để tạo tính năng vector, đó là các hashtag. Thuật toán chi-square được sử dụng để tính toán điểm quan trọng cho mỗi tính năng hoặc hashtag. Sau đó, tất cả các điểm tương ứng với các hashtag được sử dụng bởi người dùng có ảnh hưởng đã được tóm tắt. Nếu tổng là số âm, chúng ta sẽ gắn nhãn người dùng là thiên về bên trái, nếu không thì là thiên về bên phải. Guess et al. cũng suy ra khuynh hướng chính trị là một yếu tố có ý nghĩa thống kê trong sự lan truyền của tin tức giả mạo [14]. Trong số 1572 người dùng bị ảnh hưởng, chúng tôi đã có chính trị chi tiết liên kết trên 712 người dùng. Hình 4.1 cho thấy sự phân bố của thực tế và chia sẻ tin tức giả mạo giữa những người dùng khác nhau theo khuynh hướng chính trị của họ.

• Tuổi và giới tính

Theo Guess và cộng sự, một trong những yếu tố dự đoán liệu ai đó có chia sẻ thông tin giả mạo hay không tin tức có thể là tuổi của họ [14]. Hơn nữa, Shu et al. đã chứng minh rằng người dùng nữ có nhiều khả năng phát tán tin tức giả mạo hơn người dùng nam [11]. Chúng tôi đã sử dụng PyTorch việc triển khai hệ thống M3 (Đa phương thức, Đa ngôn ngữ và Đa thuộc tính) để xác định độ tuổi và giới tính của người dùng bị ảnh hưởng [23]. M3 là một hệ thống học tập cho suy luận nhân khẩu học được đào tạo trên một khối lượng lớn Bộ dữ liệu Twitter [23]. Nó có ba thuộc tính chính: (1) Đa phương thức - đầu vào có thể là hình ảnh và văn bản, (1) Đa ngôn ngữ - nó hoạt động ở 32 ngôn ngữ khác nhau



Hình 4.1: Phân bố chia sẻ tin tức thật và giả của các chính trị gia
sự sắp xếp của người dùng.

ngôn ngữ, và (3) Đa thuộc tính - nó có thể dự đoán ba thuộc tính nhân khẩu học

(giới tính, độ tuổi và trạng thái con người so với tổ chức) [23]. Đối với phân tích của chúng tôi, chúng tôi đã cung cấp

id, tên, tên màn hình, mô tả, ngôn ngữ, đường dẫn hình ảnh và kích thước đã thay đổi

tập tin hình ảnh làm đầu vào cho m3inference và độ tuổi và giới tính của người dùng kết quả

phân phối được hiển thị trong Hình 4.2. Đầu ra của chúng tôi bao gồm nhân khẩu học người dùng

ics chi tiết về tuổi tác và giới tính. Rõ ràng là công chúng lớn tuổi và phụ nữ

dễ bị thông tin sai lệch hơn.

• Phân tích ứng suất

Một loạt các nghiên cứu văn học đã chứng minh rằng tình trạng sức khỏe tâm thần của người dùng

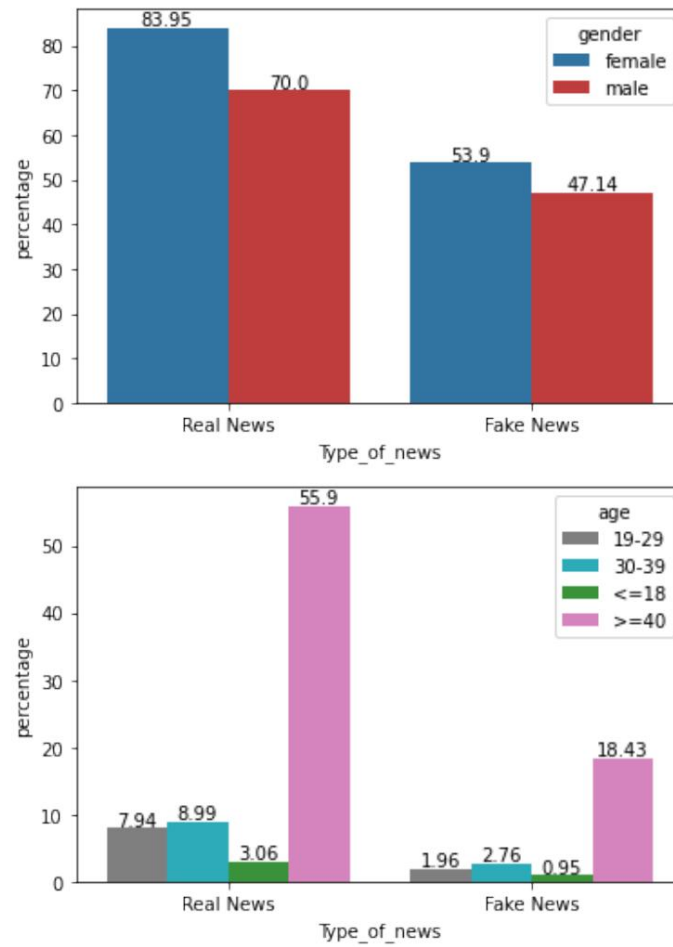
có thể dự đoán được từ các thông điệp truyền thông xã hội của họ [24]. Chúng tôi đã chuẩn bị

tập tin đầu vào bao gồm các dòng tweet được nối lại và làm sạch của từng người dùng

từ tập dữ liệu và sử dụng Pennebaker's Linguistic Inquiry and Word Count

(LIWC) cách tiếp cận để phân tích ứng suất. Là một phần của quá trình làm sạch, chúng tôi

đặt tất cả các biểu tượng cảm xúc và biểu tượng cảm xúc vào các từ dừng và dấu câu - văn bản tự do.



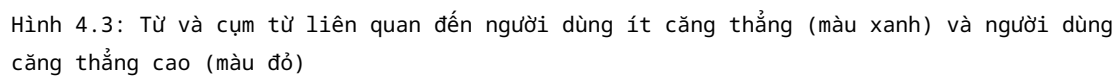
Hình 4.2: Phân bố lượng chia sẻ tin tức giả và tin tức thật theo giới tính và độ tuổi của người dùng.

Các mẫu như HTTPS, RT, #via cũng đã bị xóa.

LIWC là một chương trình phân tích văn bản minh bạch đếm các từ trong tâm lý học các danh mục có ý nghĩa về mặt khoa học [25]. Chương trình LIWC có hai tính năng—thành phần xử lý và các từ điển. Tính năng xử lý là program, mở một loạt các tệp văn bản—có thể là các bài luận, bài viết, blog, và cứ như vậy—và sau đó mỗi từ trong một tệp văn bản nhất định được so sánh với từ điển tập tin nary [26]. Kết quả thực nghiệm sử dụng LIWC chứng minh khả năng phát hiện của nó có nghĩa là trong nhiều bối cảnh thử nghiệm khác nhau. Nó giúp thể hiện sự chú ý tập trung, cảm xúc, các mối quan hệ xã hội, phong cách suy nghĩ và sự khác biệt của từng cá nhân ences. Nó cho phép người dùng xem xét kỹ hơn các tác phẩm văn học. LIWC thiết kế đã làm cho nó trở thành một yêu thích của các nhà tâm lý học, nhưng nó cũng được sử dụng trong thị trường ینگ, phân tích Twitter, chẩn đoán sức khỏe tâm thần và nhiều hơn nữa [25]. Ban đầu, phần mềm LIWC không bao gồm từ điển căng thẳng. Wang và cộng sự đã tạo ra từ điển căng thẳng theo các thủ tục và bước được Pennebaker thiết lập để đảm bảo các đặc tính tâm lý mong muốn [27]. Ngoài ra, cũng công bằng khi cho rằng từ điển căng thẳng là một từ điển phụ của từ điển cảm xúc tiêu cực trong phần mềm LIWC. Hình 4.3 mô tả các từ và cụm từ liên quan đến người dùng có mức độ căng thẳng cao và mức độ căng thẳng thấp. Nó hiển thị các từ thường được sử dụng trong số những người có mức độ căng thẳng cao và người dùng ít căng thẳng.

• Phân tích tình cảm

Phân tích tình cảm hữu ích cho nhiều vấn đề có liên quan cho các nhà nghiên cứu tương tác giữa người và máy tính. Nó cũng đã tìm thấy ứng dụng trong các lĩnh vực như xã hội học, tiếp thị và quảng cáo, tâm lý học, kinh tế, và khoa học chính trị. Phân tích tình cảm của dữ liệu Twitter đã đạt được nhiều



sự chú ý như một chủ đề nghiên cứu.

Những cảm xúc tích cực, trung lập và tiêu cực trong dòng thời gian tweet của

người dùng bị ảnh hưởng đã được tính toán bằng cách sử dụng 'VADER' (Từ điển nhận thức Valence)

cho công cụ phân tích tình cảm (Sentiment Reasoning). Phân tích tình cảm VADER

dựa vào một từ điển ánh xạ các đặc điểm từ vựng với cường độ cảm xúc được gọi là

điểm tình cảm. Điểm tình cảm của một văn bản có thể được lấy bằng cách tổng hợp

cường độ của từng từ trong văn bản. Chúng tôi đã nối tất cả các dòng tweet trên dòng thời gian

cho mỗi người dùng và các tweet đã được làm sạch được phân tích bằng VADER. Là một phần của

quá trình dọn dẹp, chúng tôi đã thay thế tất cả các biểu tượng cảm xúc và biểu tượng cảm xúc trong các từ dùng

và dấu câu - văn bản tự do. Một số mẫu, ví dụ, HTTPS, là

cũng đã bị xóa. Cực đầu ra của 1 ngụ ý tình cảm tích cực, 0 tượng trưng cho

trung lập và -1 biểu thị tâm lý tiêu cực.

- Sở thích và sự tương đồng của người dùng

Bộ dữ liệu chính cho nghiên cứu này có các cặp người có ảnh hưởng-người dùng cùng với

chi tiết về tin tức đã được chia sẻ. Để tính toán độ tương đồng cosin giữa

sở thích của người dùng và chia sẻ tin tức, chúng tôi xác định sở thích của người dùng bằng cách sử dụng

theo hai cách tiếp cận sau: (1) Lợi ích của người dùng là sự kết nối của tất cả các phần còn lại

cổ phần của người dùng bị ảnh hưởng, như đã thảo luận trong Mục 3.0.2. (2) Lợi ích của người dùng là

sự kết nối của tất cả các dòng tweet trên dòng thời gian của người dùng bị ảnh hưởng.

Đối với tính năng tính toán này, chúng tôi đã tạo một mô hình LDA trên Wikipedia bằng cách sử dụng

Gensim và trích xuất 100 chủ đề. Gensim được thiết kế để xử lý văn bản thô bằng cách sử dụng

thuật toán học máy không giám sát [28]. Các thuật toán trong Gensim, chẳng hạn như

như Word2Vec, FastText, Chỉ mục ngữ nghĩa tiềm ẩn, Phân bố Dirichlet tiềm ẩn

(LDA), v.v., khám phá cấu trúc ngữ nghĩa của tài liệu bằng cách kiểm tra trạng thái các mẫu đồng hiện diện trong một tập hợp các tài liệu đào tạo [28]. Một khi những mẫu này được tìm thấy, chúng được sử dụng để tìm ra sự tương đồng về chủ đề so với các tài liệu khác.

- Đặc điểm cảm xúc

Chúng tôi đã tính toán các đặc điểm cảm xúc bổ sung như tức giận, vui mừng, buồn bã, sợ hãi, ghê tởm, mong đợi, ngạc nhiên và tin tưởng bằng cách sử dụng Từ điển cường độ cảm xúc (NRC-EIL) [29] và cách tiếp cận được đề xuất trong [30]. Cường độ cảm xúc của NRC Từ điển là danh sách các từ tiếng Anh có điểm số cường độ thực tế cho tám những cảm xúc cơ bản (giận dữ, mong đợi, ghê tởm, sợ hãi, vui mừng, buồn bã, ngạc nhiên và tin tưởng). Đối với một từ w và cảm xúc e nhất định, điểm số nằm trong khoảng từ 0 đến 1.

- Điểm 1 cho thấy chữ w truyền tải được lượng cảm xúc cao nhất

nguồn 14

- Điểm 0 cho biết w truyền tải lượng cảm xúc thấp nhất e .

Từ điển có gần 10.000 mục từ cho tám cảm xúc. Nó bao gồm các từ phổ biến

Các thuật ngữ tiếng Anh cũng như các thuật ngữ nổi bật hơn trên nền tảng truyền thông xã hội

các hình thức, chẳng hạn như Twitter. Nó bao gồm các thuật ngữ liên quan đến cảm xúc

ở nhiều mức độ khác nhau. Các dòng tweet theo dòng thời gian được nối tiếp nhau đã được sử dụng để trích xuất điểm cường độ cảm xúc của mỗi người dùng.

- Đặc điểm hành vi

Chúng tôi đã phân tích đặc điểm thời gian và chủ đề của hành vi chia sẻ của người dùng, cho thấy cách chúng thể hiện các kiểu hành vi riêng biệt. Các đặc điểm hành vi

của người dùng được tính toán bằng cách sử dụng thông tin thời gian của các lượt chia sẻ lại và dòng thời gian tweet. Có hai tính năng được tính toán theo danh mục này. Trong

Tính năng đầu tiên, chúng tôi đã tính toán sự khác biệt thời gian giữa các bài đăng ban đêm và ban ngày bài đăng của người dùng dựa trên tổng số bài đăng của người đó. Giá trị dương cho biết rằng người dùng hoạt động chủ yếu vào ban đêm. Đối với tính năng thứ hai của chúng tôi, chúng tôi đã tính toán thời gian trung bình mà người dùng dành ra để tweet trong một khoảng thời gian nhất định tweet.

Tin tức Các tính năng liên quan:

Trong quá trình triển khai, chúng tôi đã xem xét các tính năng xác nhận những phát hiện của cả hai Shrestha và Spezzano, và Horne và Adali [31, 32]. Phong cách, tâm lý sau đây

các tính năng về mặt lý thuyết và độ phức tạp được tính toán cho cả tiêu đề và nội dung văn bản tin tức.

- Đặc điểm phong cách

Chúng tôi đã sử dụng tập hợp con các tính năng LIWC biểu diễn chức năng của văn bản, bao gồm số lượng từ (WC), số từ mỗi câu (WPS), số lượng cá nhân (Tôi, chúng ta, bạn, cô ấy/anh ấy - mỗi đặc điểm một) và đại từ vô nhân xưng, số lượng dấu chấm than (exlam), số lượng ký hiệu dấu câu (allPunc), số của trích dẫn (trích dẫn).

Về các tính năng của phần lời nói, chúng tôi đã sử dụng Python Natural Language

Bộ công cụ gắn thẻ từ loại (POS) để tính toán số lượng danh từ (NN),

danh từ riêng (NNP), đại từ nhân xưng (PRP), đại từ sở hữu (PRP),

Wh-đại từ (WP), định tố (DT), Wh-định tố (WDT), số lượng

số (CD), trạng từ (RB), động từ (VB), động từ thì quá khứ (VBD), danh động từ

hoặc động từ hiện tại phân từ (VBG), động từ quá khứ phân từ (VBN), không phải ngôi thứ 3

động từ hiện tại số ít (VBP) và động từ hiện tại số ít ngôi thứ ba (VBZ).

- Tính năng tâm lý

Tâm lý xã hội là nghiên cứu về sự tương tác năng động giữa các cá nhân

và những người xung quanh họ. Tâm lý học đóng vai trò quan trọng trong lĩnh vực này

của tiếp thị truyền thông xã hội. Người ta cần khai thác cảm xúc để phát triển

mối quan hệ khách hàng lâu dài. Khoa học tâm lý xã hội đã ra đời

sự tồn tại khi các nhà khoa học lần đầu tiên bắt đầu đo lường chính thức những suy nghĩ, cảm xúc, và hành vi của con người.

Chúng tôi đã tính toán các số liệu tình cảm tích cực (tích cực) và tiêu cực (tiêu cực) bằng cách sử dụng

công cụ LIWC. Như Shrestha và Spezzano, [32] và Ghanem et al., [33] đã nêu lại

gần đây đã chỉ ra rằng cảm xúc đóng vai trò quan trọng trong việc đánh lừa người đọc và có thể

thành công được sử dụng để phát hiện thông tin sai lệch. Ngoài tình cảm

số liệu, chúng tôi đã tính toán các đặc điểm cảm xúc, chẳng hạn như tức giận, vui mừng, buồn bã, sợ hãi, đau khổ

cơn gió, sự mong đợi, sự ngạc nhiên và sự tin tưởng bằng cách sử dụng Từ điển cường độ cảm xúc

(NRC-EIL) [29] và cách tiếp cận được đề xuất trong [30]. Từ điển cảm xúc NRC

là danh sách các từ tiếng Anh và mối liên hệ của chúng với tám cảm xúc cơ bản (tức giận,

sợ hãi, mong đợi, tin tưởng, ngạc nhiên, buồn bã, vui mừng và ghê tởm) và hai cảm xúc

(tiêu cực và tích cực). Chúng tôi đã tính toán những điểm số này cho cả văn bản và tiêu đề của

tin tức khác nhau.

- Tính năng phức tạp

SMOG - Độ phức tạp của văn bản trong xử lý ngôn ngữ tự nhiên phụ thuộc vào cách

người đọc có thể dễ dàng đọc và hiểu một văn bản. Chúng tôi đã sử dụng Biện pháp đơn giản

của biện pháp đo lường khả năng đọc Gobbledygook Index (SMOG) như một tính năng phức tạp trong phân tích của chúng tôi [34]. Công thức khả năng đọc ước tính số năm giáo dục người ta cần hiểu một tác phẩm văn học. McLaughlin đã tạo ra công thức này như một sự cải tiến so với các công thức khả năng đọc khác [34]. Điểm số cao hơn của những công thức này Các biện pháp dễ đọc cho thấy văn bản dễ đọc hơn.

Sự đa dạng từ vựng - Đây là phép đo về số lượng từ vựng khác nhau có trong một văn bản. Từ vựng là những từ như danh từ, tính từ, động từ, và các trạng từ truyền đạt ý nghĩa trong văn bản.

Tỷ lệ loại-mã thông báo (TTR) - Tỷ lệ loại-mã thông báo (TTR) là tổng số Các từ DUY NHẤT (loại) chia cho tổng số từ (mã thông báo) trong một phân đoạn ngôn ngữ nhất định. Tỷ lệ TTR càng gần 1 thì càng lớn sự phong phú về mặt từ vựng của nội dung.

Độ dài từ trung bình (avg wlen) - Độ dài trung bình của các từ (số lượng ký tự-điển viên).

Các tính năng liên quan đến mạng:

- Mối quan hệ yếu và mạnh

Theo Ma et al., sức mạnh gắn kết được nhận thức trong các mạng xã hội trực tuyến là tích cực có liên quan tích cực đến ý định chia sẻ tin tức trên phương tiện truyền thông xã hội [13]. M. Granovetter cho thấy mối quan hệ chặt chẽ là bạn bè và mối quan hệ yếu là người quen biết [35]. Bài báo nói về mối quan hệ giữa các cá nhân giữa các nhóm khác nhau của mọi người và cách họ giữ mạng lưới của họ lại với nhau. Theo Granovetter, thông tin lan truyền nhanh hơn giữa những người có mối quan hệ chặt chẽ. Trong khi mối quan hệ chặt chẽ

chỉ ra số lượng lớn cổ phần giữa hai người trong một mạng lưới, mối liên kết yếu

mô tả ít lượt chia sẻ hơn giữa người có sức ảnh hưởng và người theo dõi.

Mặc dù có thể có những chiến lược khác nhau để xác định sức mạnh liên kết giữa

cặp người có ảnh hưởng và người dùng bị ảnh hưởng, nghiên cứu tập trung vào những điều sau đây

hai định nghĩa.

Quan điểm của người nhận - Đối với một cặp người có ảnh hưởng-người theo dõi nhất định trong tập dữ liệu,

tỷ lệ phần trăm các tweet được người dùng có ảnh hưởng chia sẻ lại trên các tweet của người dùng có ảnh hưởng

đã được tính toán.

Phân tích dựa trên thời gian - Điều này được tính là thời gian trung bình được thực hiện bởi một

người dùng bị ảnh hưởng chia sẻ từ người có sức ảnh hưởng.

- Tỷ lệ người theo dõi trên Twitter (TFF)

Điều này được tính toán như là tỷ lệ giữa số người theo dõi và số người đang theo dõi. Dữ liệu từ

FakeNewsNet có thông tin chi tiết về người dùng và những người theo dõi họ. Công thức

để tính toán TFF được hiển thị bên dưới. Thêm 1 giữ nguyên tỷ lệ

từ giảm xuống vô cực hoặc bằng không.

$$TFF = \frac{\#Follower + 1}{\#Following + 1}$$

Trong khi tỷ lệ 2.0 trở lên cho thấy bạn là người nổi tiếng và được nhiều người yêu mến.

muốn nghe những gì bạn phải nói, tỷ lệ khoảng 1.0 có nghĩa là bạn

được tôn trọng giữa các đồng nghiệp của bạn. TFF được đề xuất như một tính năng mạng quan trọng

bởi Shu và cộng sự [11].

- Pagerank và mức độ trung tâm

PageRank tính đến hướng liên kết. Nó có thể giúp khám phá ra những người có ảnh hưởng hoặc quan trọng các nút quan trọng có phạm vi mở rộng vượt ra ngoài các kết nối trực tiếp của chúng. Mỗi nút trong mạng được gán điểm dựa trên số lượng liên kết đến của nó (in-bậc). Phương trình đệ quy cho PageRank được đưa ra bởi,

$$PR(p_i) = \frac{1}{N} + \sum_{p_j \in M(p_i)} \frac{PR(p_j)}{L(p_j)}$$

trong đó $M(p_i)$ là tập hợp các nút trỏ tới p_i và $L(p_j)$ là số lượng các nút p_j trỏ tới.

Độ trung tâm của cấp độ cho biết tầm quan trọng tương đối của một nút trong mạng.

Nhìn chung, đối với một nút x nhất định, nó được tính là tỷ lệ giữa số lượng các nút được kết nối với nút x và tổng số tất cả các nút trong mạng

(giảm đi một). Chúng tôi đã tính toán độ trung tâm trong và ngoài bằng cách sử dụng thư viện, `networkx`. Bộ dữ liệu được sử dụng từ FakeNewsNet chứa thông tin chi tiết về người dùng và những người theo dõi họ.

Vì nghiên cứu của chúng tôi sử dụng lý thuyết DOI nên nó yêu cầu các tính năng từ mỗi danh mục tin tức, người dùng và mạng. Điều này đã làm tăng số lượng tính năng trong nghiên cứu của chúng tôi.

Các thuật toán học máy có khả năng điều tra nhiều tính năng so với phân tích thủ công. Do đó, nó phù hợp với mục đích.

4.1 Phương pháp

Đối với phần đầu tiên của câu hỏi nghiên cứu được đề cập trong Phần 1, các bộ phân loại chẳng hạn như Hồi quy Logistic, Máy vectơ hỗ trợ (SVM), Rừng ngẫu nhiên, XG-Boost và Extra Trees Classifier.

4.1.1 Phân tích sự sống còn

Như đã đề cập trong Chương 1, chúng tôi cũng sẽ dự đoán thời gian chia sẻ nội dung theo người dùng, sử dụng Rừng sinh tồn ngẫu nhiên. Gói Rừng sinh tồn ngẫu nhiên cung cấp một bản triển khai python của phương pháp dự đoán sự sống còn ban đầu được xuất bản bởi Ishwaran và cộng sự [36].

Tổng quan

Phân tích sinh tồn là một lĩnh vực thống kê trong đó mục tiêu là phân tích và mô hình hóa dữ liệu trong đó kết quả là thời gian cho đến khi sự kiện quan tâm xảy ra [37]. Nói chung, các phương pháp phân tích sự sống còn có thể được phân loại thành hai loại chính: thống kê phương pháp và phương pháp dựa trên máy học. Các phương pháp thống kê chia sẻ một mục tiêu với các phương pháp học máy là cả hai đều được kỳ vọng sẽ đưa ra dự đoán của thời gian sống sót và ước tính xác suất sống sót ở thời điểm sống sót ước tính thời gian. Các phương pháp học máy thường được áp dụng cho các vấn đề có chiều cao, trong khi các phương pháp thống kê thường được phát triển để xử lý dữ liệu có chiều thấp.

Tùy thuộc vào các giả định được đưa ra và cách các tham số được sử dụng trong mô hình, các phương pháp thống kê truyền thống có thể được chia thành ba loại: (i) mô hình phi tham số, (ii) mô hình bán tham số và (iii) mô hình tham số.

Các thuật toán học máy như cây sống sót, phương pháp Bayesian, mạng nơ-ron các tác phẩm và máy hỗ trợ vector được đưa vào một nhánh riêng biệt. Một số phương pháp học máy tiên tiến, bao gồm học tập tổng hợp, học tập chủ động, phương pháp học chuyển giao và phương pháp học đa nhiệm cũng được bao gồm trong danh mục này. Đối tượng quan tâm chính là hàm sống sót, thường được ký hiệu là S , được định nghĩa là

$$S(t) = \Pr(T > t)$$

trong đó t là thời gian, T là biến ngẫu nhiên biểu thị thời gian của một sự kiện và " \Pr " là viết tắt của xác suất. Nghĩa là, hàm sống sót là xác suất mà thời gian của một sự kiện xảy ra muộn hơn một thời điểm t xác định.

4.1.2 Kiểm duyệt

Kiểm duyệt là phổ biến trong phân tích sinh tồn. Đây là một dạng vấn đề dữ liệu bị thiếu trong thời gian nào sự kiện không được quan sát vì,

- Về việc chấm dứt nghiên cứu trước khi tất cả các đối tượng được tuyển dụng đã trình bày sự kiện quan tâm hoặc
- Đối tượng đã rời khỏi nghiên cứu trước khi trải qua sự kiện.

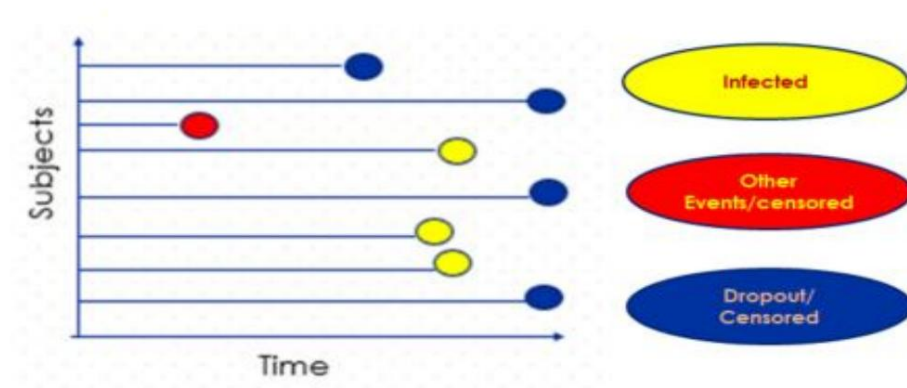
Các loại

Kiểm duyệt bên trái - Nếu sự kiện quan tâm đã xảy ra trước khi chủ đề được được đưa vào nghiên cứu nhưng không biết khi nào nó xảy ra, dữ liệu được cho là bên trái bị kiểm duyệt.

Kiểm duyệt đúng - Nếu chỉ biết giới hạn dưới l cho thời gian sự kiện thực T như vậy

T đó, điều này được gọi là kiểm duyệt đúng. Kiểm duyệt đúng sẽ xảy ra, ví dụ, đối với những đối tượng có ngày sinh được biết đến nhưng vẫn còn sống khi họ mất tích theo dõi hoặc khi nghiên cứu kết thúc. Chúng tôi thường gặp dữ liệu bị kiểm duyệt bên phải và Hình 4.4 cho thấy biểu diễn bằng hình ảnh với thời gian học so với đối tượng học.

Kiểm duyệt khoảng thời gian - Khi chỉ có thể nói rằng sự kiện xảy ra giữa hai quan sát hoặc kiểm tra, đây là kiểm duyệt theo khoảng thời gian.



Hình 4.4: Kiểm duyệt bên phải

Rừng sinh tồn ngẫu nhiên

Phương pháp này được sử dụng với dữ liệu sống sót được kiểm duyệt bên phải [36]. Kiểm duyệt bên phải xảy ra khi một đối tượng rời khỏi nghiên cứu trước khi một sự kiện xảy ra, hoặc nghiên cứu kết thúc trước khi sự kiện đã xảy ra. Một khu rừng sống sót ngẫu nhiên bao gồm các cây sống sót ngẫu nhiên. Sử dụng mẫu bootstrap độc lập, mỗi cây được trồng bằng cách chọn ngẫu nhiên một tập hợp con các biến tại mỗi nút và sau đó chia tách nút bằng cách sử dụng tiêu chí sống sót liên quan đến thời gian sống sót và thông tin trạng thái kiểm duyệt. Cây được coi là đã trưởng thành hoàn toàn khi mỗi nút đầu cuối có không ít hơn $d_0 > 0$ ca tử vong duy nhất. Ước tính hàm nguy cơ tích lũy (CHF) cho một trường hợp là ước lượng Nelson-Aalen cho

nút cuối của trường hợp. Tập hợp là trung bình của các CHF này. Bởi vì cây là được phát triển từ dữ liệu trong túi, một tập hợp ngoài túi (OOB) có thể được tính toán bằng cách loại bỏ Các trường hợp OOB xuống cây sống sót trong túi của họ và tính trung bình. Giá trị dự đoán cho một trường hợp sử dụng nhóm OOB không sử dụng thông tin sinh tồn cho trường hợp đó và, do đó, nó có thể được sử dụng để ước tính gần như không thiên vị về lỗi dự đoán. Từ điều này, các biện pháp hữu ích khác có thể được rút ra, chẳng hạn như các giá trị tầm quan trọng biến đổi cho lọc và lựa chọn biến.



Hình 4.5: Cây sống sót ngẫu nhiên

Chức năng sống sót và ước tính thời gian

Cho một trường hợp mới được mô tả bởi vectơ đặc trưng X_i , ước tính phân tích sự sống còn một hàm sinh tồn S_i cung cấp xác suất rằng sự kiện cho trường hợp i sẽ xảy ra sau thời điểm t , tức là $S_i(t) = P(T_i \geq t)$.

Giả sử $i = (u, v)$ là cặp người có ảnh hưởng - người theo dõi. Xem xét trường hợp người dùng u là theo sau là người dùng v và người dùng v được dự đoán sẽ chia sẻ nội dung từ người dùng u trong khoảng thời gian $[t_a, t_b]$. Thời gian khi sự kiện quan tâm xảy ra đối với trường hợp $i = (u, v)$ được biểu thị bằng $T_i(T_i \in [t_a, t_b])$. Chúng tôi cũng giả sử khoảng thời gian $[t_a, t_b]$ là chia thành k khoảng thời gian, ví dụ, ngày, tuần, tháng, v.v. Từ phân tích sinh tồn mô hình, chúng ta sẽ có các xác suất sau: $S_i(t_a) = P(T_i > t_a)$, $S_i(t_a + k) = P(T_i > t_a + k)$, $S_i(t_b) = P(T_i > t_b)$. Xác suất chia sẻ sẽ xảy ra

trong khoảng thời gian $[t_a + (h-1)k, t_a + hk)$ được đưa ra bởi,

$$P_r(t_a + (h-1)k \leq T_i < t_a + hk) = P_r(T_i \geq t_a + (h-1)k) - P_r(T_i \geq t_a + hk)$$

Từ hàm sống sót, hàm mật độ xác suất là sự kiện sẽ

xảy ra trong khoảng thời gian cho trước $[t_a + (h-1)k, t_a + hk)$. Giả sử x biểu thị tất cả các khoảng thời gian trong

hàm mật độ xác suất và thời gian là giá trị mong đợi, $E(x)$, được ước tính

BẰNG,

$$E(x) = \int_{t_a}^{\infty} x \cdot (P_r(t_a + (h-1)k \leq T_i < t_a + hk))$$

$$E(x) = \sum_{x=t_a}^{\infty} x \cdot (P_r(t_a + (h-1)k \leq T_i < t_a + hk))$$

CHƯƠNG 5:

THÍ NGHIỆM VÀ KẾT QUẢ

5.1 Thiết lập thử nghiệm

Chúng tôi đã thử nghiệm các tính năng của mình bằng cách sử dụng phân loại nhị phân để xác định xem người dùng có chia sẻ hay không hoặc không phải là một tin tức nhất định từ người có ảnh hưởng của anh ấy/cô ấy trên nhiều máy học khác nhau các thuật toán cụ thể là Hồi quy logistic, Máy vectơ hỗ trợ (SVM), Ngẫu nhiên Forest, XGBoost và Extra Trees Classifier. Vì tập dữ liệu rất không cân bằng được nâng cấp (nhãn 1: 3.144, nhãn 0: 14.936), chúng tôi đã sử dụng trọng số lớp để xử lý nó. Chúng tôi cũng sử dụng phương pháp quy kết đơn giản bằng cách sử dụng giá trị trung bình để có được quan sát viên bị thiếu. Để đánh giá hiệu suất, chúng tôi đã xem xét Diện tích bên dưới Máy thu Đặc điểm hoạt động (AUROC) và Độ chính xác trung bình, rất phù hợp cho dữ liệu không cân bằng và thực hiện xác thực chéo 10 lần. Bộ phân loại XGBoost vượt trội hơn trong số họ về việc chia sẻ tin tức giả và thật. Bộ phân loại được đưa ra một av- độ chính xác của erage là 95,23 với AUROC là 97,39 để chia sẻ tin tức thực sự. Đối với tin tức giả chia sẻ, nó báo cáo AUROC và độ chính xác trung bình lần lượt là 97,34 và 88,43, như thể hiện trong Bảng 5.1 và 5.2. XGBoost là một máy tổng hợp dựa trên cây quyết định thuật toán học tập sử dụng khuôn khổ tăng cường độ dốc.

5.1.1 So sánh với các đường cơ sở

Chúng tôi đã so sánh kết quả của bộ phân loại với các đường cơ sở trong Chương 2: 1) Độc lập Mô hình thác đổ phụ thuộc và 2) Mô hình ngưỡng tuyến tính. Thuật toán phân loại trong học máy sử dụng dữ liệu đào tạo đầu vào để dự đoán khả năng xảy ra tiếp theo dữ liệu sẽ rơi vào một trong những loại được xác định trước. Khả năng là xác suất giá trị khả năng. Chúng tôi đã sử dụng Phân phối Bernoulli và Chỉ số Jaccard để xác định điều này xác suất cho mô hình thác nước độc lập [38]. Không giống như các bộ phân loại dựa trên ML, xác suất lan truyền của đường cơ sở không phụ thuộc vào tin tức hoặc tính năng của người dùng.

Phân phối Bernoulli

Trong mô hình này, nút được kích hoạt 'v' ảnh hưởng đến nút không hoạt động 'u' bằng cách sử dụng một nút cố định xác suất hoặc ngưỡng. Một nỗ lực thành công là khi 'u' được kích hoạt. Mỗi nỗ lực, có liên quan đến một số hành động, có thể được xem như một thử nghiệm Bernoulli. xác suất thành công là tỷ lệ số lần thử thành công trên tổng số số lần thử nghiệm [38]. Do đó, xác suất ảnh hưởng của 'v' đối với 'u' bằng cách sử dụng Tối đa Ước tính khả năng (MLE) được ước tính như sau:

$$p_{v,u} = \frac{A_{v2u}}{A_v}$$

Chỉ số Jaccard

Nó được định nghĩa là kích thước của giao điểm chia cho kích thước của hợp của mẫu đang nghiên cứu. Goyal và cộng sự đã điều chỉnh Chỉ số Jaccard như sau [38],

$$p_{v,u} = \frac{A_{v2u}}{A_u \mid}$$

Trong Mô hình ngưỡng tuyến tính, mỗi cạnh $e = (u, v)$ được liên kết với một trọng số bu, v , như đã giải thích trong Phần 2. Ngoài ra, có một xác suất lan truyền cần thiết cho sự lan truyền thông tin. Trong nghiên cứu này, chúng tôi đã sử dụng hai phương pháp để tính toán xác suất. Trong cách tiếp cận đầu tiên, nghịch đảo của mức độ ảnh hưởng của người dùng được nhân với số lượng người có ảnh hưởng như xác suất kích hoạt [39]. Thứ hai, cách tiếp cận lấy các giá trị từ một tập hợp $[0, 1, 0, 01, 0, 001]$ và nó được nhân với số lượng người có ảnh hưởng đến từng người dùng trong tập dữ liệu chính [39]. Sau khi xác suất được tính toán cho cả mô hình Ngưỡng tuyến tính và Độc lập Cascade, chúng tôi đã sử dụng 10 - xác thực chéo và đánh giá mô hình bằng AUROC và độ chính xác trung bình. Chúng tôi cũng sử dụng Kỹ thuật lấy mẫu quá mức thiếu số tổng hợp (SMOTE) để xử lý dữ liệu mất cân bằng.

Kết quả của các đường cơ sở cũng được báo cáo trên Bảng 5.1 và 5.2. Trong số hai mô hình cơ sở, Mô hình thác nước độc lập (ICM) vượt trội hơn với AUROC của 67,67 và độ chính xác trung bình là 56,73 đối với việc chia sẻ tin tức giả mạo. Đối với việc chia sẻ tin tức thực sự, ICM đưa ra AUROC là 64,42 và độ chính xác trung bình là 48,54. Do đó, chúng ta có thể kết luận rằng mô hình của chúng tôi hoạt động tốt hơn các mô hình cơ sở. Không giống như ML - dựa trên các bộ phân loại, các đường cơ sở thiếu nhiều tính năng từ người dùng, tin tức, mạng và do đó giải thích AUROC thấp và độ chính xác trung bình.

5.1.2 Phân tích việc chia sẻ tin tức thật và giả

Các số liệu giúp giải thích dự báo xác suất cho lớp nhị phân

Các vấn đề về kích thước là đường cong ROC và đường cong Precision-Recall. Đường cong ROC là thích hợp khi các quan sát được cân bằng giữa mỗi lớp, trong khi độ chính xác

đường cong thu hồi thích hợp cho các tập dữ liệu mất cân bằng. Shrestha và Spezzano, đã sử dụng

Bảng 5.1: Kết quả phân loại: Kết quả phân loại cho Fake News Sharing

Tên mẫu	Độ chính xác trung bình của AUROC	
Hồi quy logistic	93,78	65,22
SVM	89,08	36,58
Rừng ngẫu nhiên	97,86	85,11
Cây bổ sung	96,82	76,75
XGBoost	97,34	88,43
Baselines		
Mô hình thác nước độc lập	67,70	63,78
Mô hình ngưỡng tuyến tính		87,45

Bảng 5.2: Kết quả phân loại: Kết quả phân loại cho Real News Sharing

Tên mẫu	Độ chính xác trung bình của AUROC	
Hồi quy logistic	92,39	84,15
SVM	87,05	66,91
Rừng ngẫu nhiên	97,38	94,61
Cây bổ sung	95,26	85,94
Tăng cường XG	97,39	95,23
Đường cơ sở		
Mô hình thác nước độc lập	64,48	59,35
Mô hình ngưỡng tuyến tính		63,77

cùng một tập dữ liệu từ Politifact để phát hiện tin tức giả mạo. Chúng tôi đã sử dụng tin tức

các tính năng từ Shrestha và Spezzano, và nhận được AUROC gần 93% (sử dụng

chỉ các tính năng tin tức) để dự đoán việc chia sẻ tin tức giả và thật [32].

Các số liệu trong 5.1 giải thích hiệu suất của bộ phân loại cho từng nhóm tính năng,

chẳng hạn như người dùng, mạng và tin tức. Trong khi các tính năng tin tức đóng góp 93%

ra quyết định, điều này có thể được thúc đẩy bằng cách sử dụng mạng lưới và các tính năng người dùng của chúng tôi. Chúng tôi đã

tổng thể tăng 4% trong AUROC như đã báo cáo trong Bảng 5.1 và 5.2.

hình ảnh trực quan của nghiên cứu cắt bỏ đặc điểm được trình bày trong Hình 5.2, 5.3, 5.4 và 5.5.

tiêu đề tin tức và các tính năng văn bản luôn mang lại điểm AUROC tốt cho mỗi bộ phân loại.

Hình 5.3 cho thấy độ mạnh của liên kết là một tính năng quan trọng đối với loại mạng.

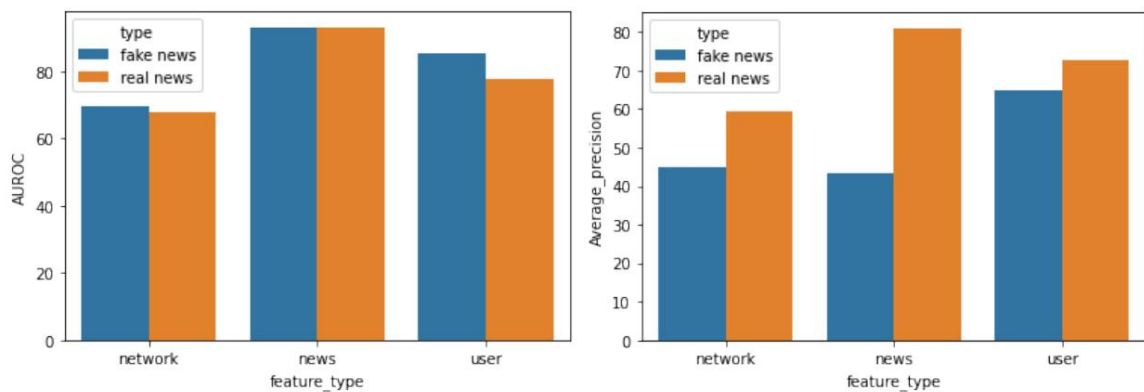
Ngoài ra, theo danh mục người dùng, độ tương đồng cosin giữa sở thích của người dùng và sở thích chung

tin tức là một tính năng chính cùng với các tính năng rõ ràng. Độ chính xác trung bình chỉ ra rằng

mỗi lần một bộ phân loại dự đoán một lớp, thì tỷ lệ phần trăm mà bộ phân loại dự đoán là bao nhiêu?

đúng. Độ chính xác trung bình thấp khi chia sẻ tin tức giả là do mẫu của nó nhỏ

kích cỡ.



Hình 5.1: Đóng góp từ mỗi Nhóm tính năng

Những hiểu biết sâu sắc về tính cách từ IBM Watson đã được nghiên cứu thêm bằng cách sử dụng ngẫu nhiên



Hình 5.2: Kết quả loại bỏ tính năng (Tin tức)

tầm quan trọng của tính năng dựa trên rừng và điểm số được hiển thị trong Hình 5.6. Tính năng tầm quan trọng là một kỹ thuật gán điểm quan trọng cho các tính năng. Điểm chỉ ra tầm quan trọng tương đối của từng tính năng khi đưa ra dự đoán. Nó có thể được sử dụng để cải thiện mô hình dự đoán. Năm tính năng hàng đầu cho cả giả và thật chia sẻ tin tức được chọn là quan trọng cho cuộc thảo luận này. 'Cần' mô tả các khía cạnh của Tweet có thể sẽ gây được tiếng vang với người dùng và 'Giá trị' cho biết các yếu tố thúc đẩy ảnh hưởng đến người dùng chia sẻ tin tức đó. Đầu trang năm tính năng liên quan đến 'Cần', cho thấy người dùng chia sẻ tin tức thực sự, hiển thị các đặc điểm như lý tưởng và tình yêu. Mặt khác, các tính năng 'cần' quan trọng được quan sát thấy đối với người dùng chia sẻ tin tức giả bao gồm sự tò mò, gần gũi, cấu trúc và sự hòa hợp. Trong số



Hình 5.3: Kết quả loại bỏ tính năng (Mạng)

'giá trị', sự tự nâng cao và sự cởi mở để thay đổi có tầm quan trọng cao trong tin tức giả mạo chia sẻ. Để chia sẻ tin tức thực sự, tự vượt qua, bảo tồn, cởi mở với sự thay đổi là quan trọng. Phân tích này cung cấp thông tin chi tiết về đặc điểm của người dùng và thiết lập một số mối tương quan giữa tính cách và thói quen chia sẻ tin tức.

Nghiên cứu này kết hợp các ảnh hưởng đa cấp của người dùng, tin tức và mạng lưới thuộc tính. Sau khi sàng lọc các tính năng tốt nhất, chúng tôi đã phân tích các tính năng đó trên

cùng một bộ phân loại. Hình 5.8 cho thấy chế độ xem đơn giản hóa của tất cả các tính năng chính có AUROC tối đa và độ chính xác trung bình. Rõ ràng là sức mạnh của mối liên kết, các tính năng tin tức và các thuộc tính tương đồng rõ ràng và người dùng, đóng góp nhiều hơn cho mức độ tách biệt. Chúng tôi tiếp tục thực hiện phân loại bằng cách sử dụng các tính năng chính và kết quả được trình bày trong Bảng 5.4 và 5.3. Bộ phân loại XGBoost đã vượt trội hơn trong cả việc chia sẻ tin tức thật và giả. Sử dụng các tính năng chính, trình phân loại XGBoost được AUROC là 97,24 cho việc chia sẻ tin tức giả mạo và đạt được AUROC là 97,53 để chia sẻ tin tức thực sự. Để quan sát các yếu tố phân biệt giữa tin giả và chia sẻ tin tức thực sự, chúng tôi đã tiến hành một nghiên cứu về tầm quan trọng của các tính năng bằng cách sử dụng rừng ngẫu nhiên trên các tính năng chính. Kết quả được hiển thị trên Hình 5.7. Trong số các tính năng hàng đầu, việc chia sẻ tin tức giả mạo rất coi trọng số lượng danh từ riêng, trạng từ, danh từ sở hữu, số từ trong mỗi câu và cảm xúc tiêu cực. Và chia sẻ tin tức thực sự phụ thuộc nhiều hơn vào sức mạnh liên kết, sự tương đồng, dấu câu, số từ trên một câu và cảm xúc tiêu cực.

5.1.3 Thiết lập thử nghiệm cho phân tích sinh tồn

Như đã đề cập trong Phần 4.0.2, chúng tôi đã sử dụng Rừng sóng sót ngẫu nhiên (RSF) để dự đoán thời gian cho sự kiện nghiên cứu của chúng tôi. Các triển khai của mô hình có sẵn trong Python [40]. Thời gian đào tạo cho gói này khá cao (hơn 48 giờ), do đó chúng tôi đã sử dụng khoảng 800 trường hợp ngẫu nhiên từ tập dữ liệu chính cho cả hai chia sẻ tin tức thật và giả. Khi các thuật toán học máy có xu hướng tăng độ chính xác bằng cách giảm lỗi, chúng không giải quyết được vấn đề phân phối lớp hoặc tập dữ liệu mất cân bằng. Để xử lý dữ liệu không cân bằng, lấy mẫu quá mức bằng cách sử dụng Synthetic Minority Oversampling Kỹ thuật (SMOTE) đã được sử dụng và thuật toán được chạy tiếp trên phép thử chéo 5 lần

xác thực. Bộ dữ liệu có tất cả các tính năng được tính toán cho người dùng, tin tức và mạng các danh mục. Chi tiết của tập dữ liệu sau khi lấy mẫu quá mức được hiển thị trong Bảng 5.5 và 5.6. Thời gian sự kiện là thời gian người dùng phản hồi một tin tức nhất định.

Bộ dữ liệu đã được kiểm duyệt bên phải và gán giá trị -1 cho các trường hợp bị kiểm duyệt.

Đối với hồi quy, chúng tôi đã xem xét hồi quy Ridge và Lasso. Vì hồi quy không thể xử lý các trường hợp bị kiểm duyệt, chúng tôi ước tính thời gian xảy ra của các trường hợp đó với giá trị thời gian $t_c \gg t_b$.

Để so sánh hiệu suất của RSF với các giá trị cơ sở, chúng tôi đã xem xét hai số liệu, cụ thể là chỉ số phù hợp và Sai số tuyệt đối trung bình (MAE). Bảng 5.7 và 5.8 cho thấy kết quả phân tích của chúng tôi. Sự khác biệt giữa dự đoán và thực tế thời gian chia sẻ nội dung trên Twitter đã cấu thành MAE trong mô hình hồi quy. Để so sánh chuỗi thời gian của số lượng đơn vị thực tế và dự đoán đã trải qua sự kiện tại mỗi thời điểm t trong khu rừng sống sót ngẫu nhiên. Chúng tôi đã tính toán thực tế hàm mật độ của dữ liệu sử dụng ước lượng Kaplan-Meier và so sánh nó với trung bình của tất cả các hàm mật độ dự đoán. Từ phân tích, trong khi các đường cơ sở vượt trội hơn phương pháp đề xuất trong việc chia sẻ tin tức thực tế, RSF bằng điểm c-index với các đường cơ sở để chia sẻ tin tức giả mạo. Chỉ số phù hợp hoặc chỉ số c phụ thuộc về thứ tự của các trường hợp trong tập dữ liệu. Chỉ số thống nhất cho tin tức thực tế chia sẻ phát hiện thấp hơn một chút so với đường cơ sở và chỉ số phù hợp của Rừng sống sót ngẫu nhiên (RSF) có thể so sánh với đường cơ sở để chia sẻ tin tức giả mạo. Hơn nữa, khi so sánh với các mô hình cơ sở hồi quy, Giá trị trung bình tuyệt đối Lỗi (MAE) trong RSF ít hơn đáng kể khi chia sẻ cả tin tức thật và tin tức giả.

Bảng 5.3: Các tính năng chính: Kết quả phân loại cho việc chia sẻ tin tức giả mạo

Tên mô hình AUROC	Độ chính xác	trung bình
Rừng ngẫu nhiên 96,90	85,28	
Cây bổ sung	96,59	81,90
Tăng cường XG	97,24	87,43

Bảng 5.4: Các tính năng chính: Kết quả phân loại cho Real News Sharing

Tên mô hình AUROC	Độ chính xác	trung bình
Rừng ngẫu nhiên 97.08	94.24	
Cây bổ sung	95,92	90,14
Tăng cường XG	97,53	95,24

Bảng 5.5: Bộ dữ liệu - Chia sẻ tin tức giả

	Chính trị hiện vật
Số lượng người dùng	80
Người dùng bị ảnh hưởng	352
Không ảnh hưởng đến người dùng	352

Bảng 5.6: Bộ dữ liệu - Chia sẻ tin tức thực tế

	Chính trị hiện vật
Số lượng người dùng	80
Người dùng bị ảnh hưởng	399
Không ảnh hưởng đến người dùng	399

Bảng 5.7: Kết quả hồi quy: Kết quả hồi quy cho việc chia sẻ tin tức thực tế

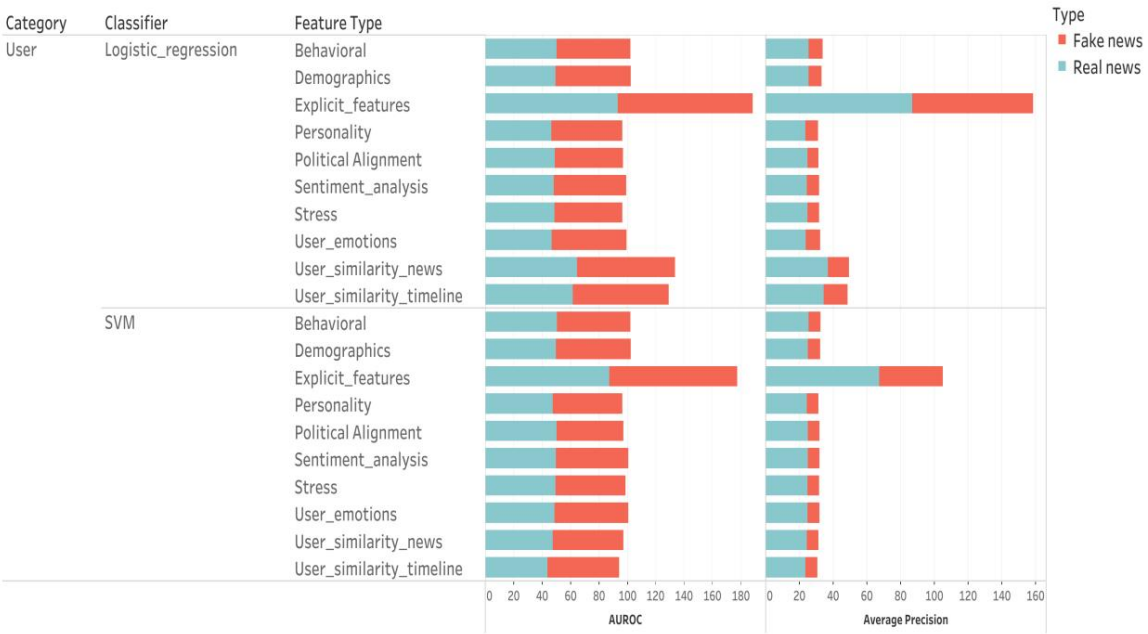
Tên mẫu	C-Index Sai	số tuyệt đối trung bình
Rừng sinh tồn ngẫu nhiên	0,68	1.29
Đường cơ sở		
Hồi quy Ridge	0,768	53605900
hồi quy Lasso	0,77	53547000

Bảng 5.8: Kết quả hồi quy: Kết quả hồi quy cho việc chia sẻ tin tức giả mạo

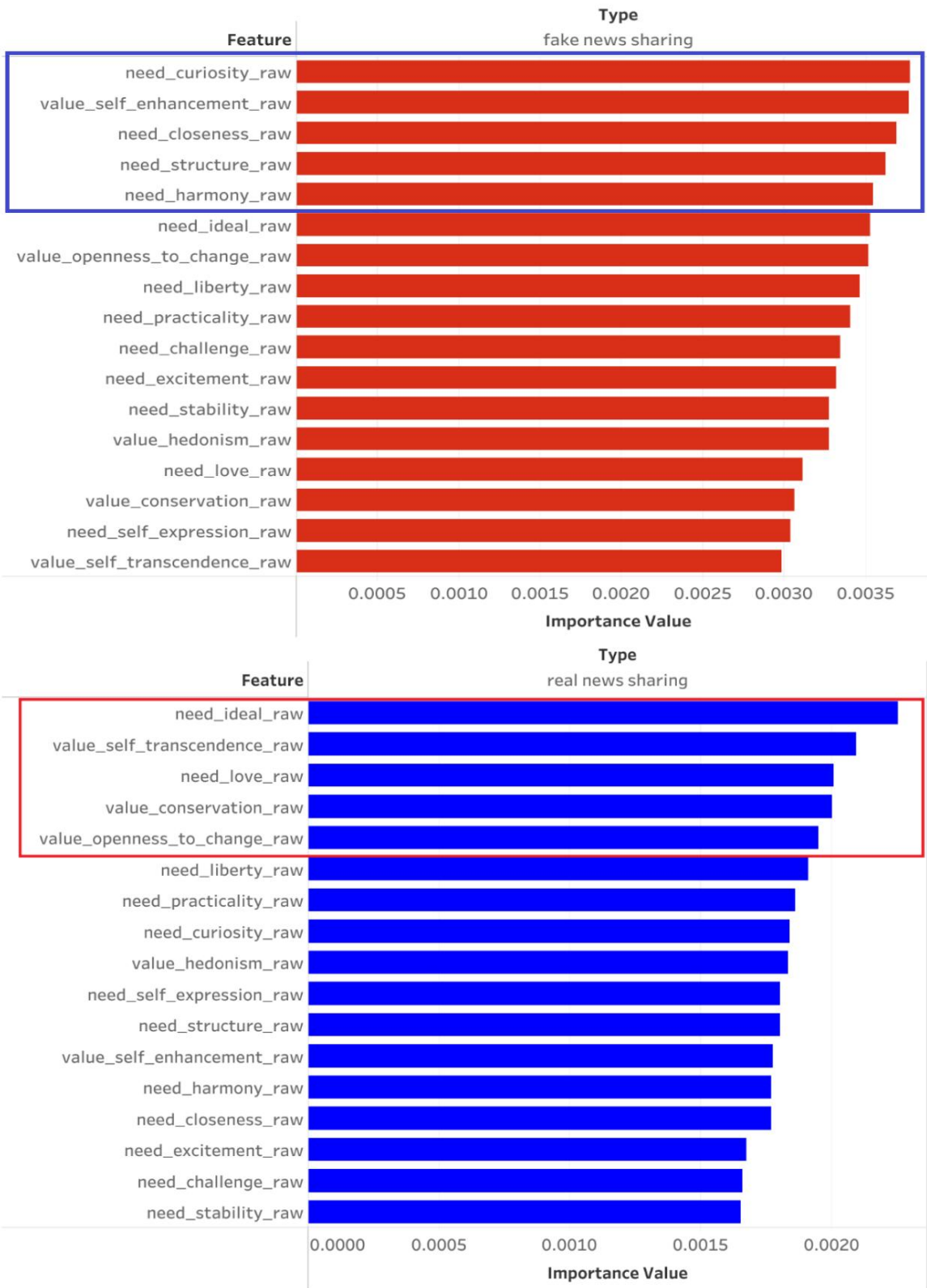
Tên mẫu	C-Index Sai	số tuyệt đối trung bình
Rừng sinh tồn ngẫu nhiên	0,81	1,25
Đường cơ sở		
Hồi quy Ridge	0,81787	40180300
hồi quy Lasso	0,80855	41328300



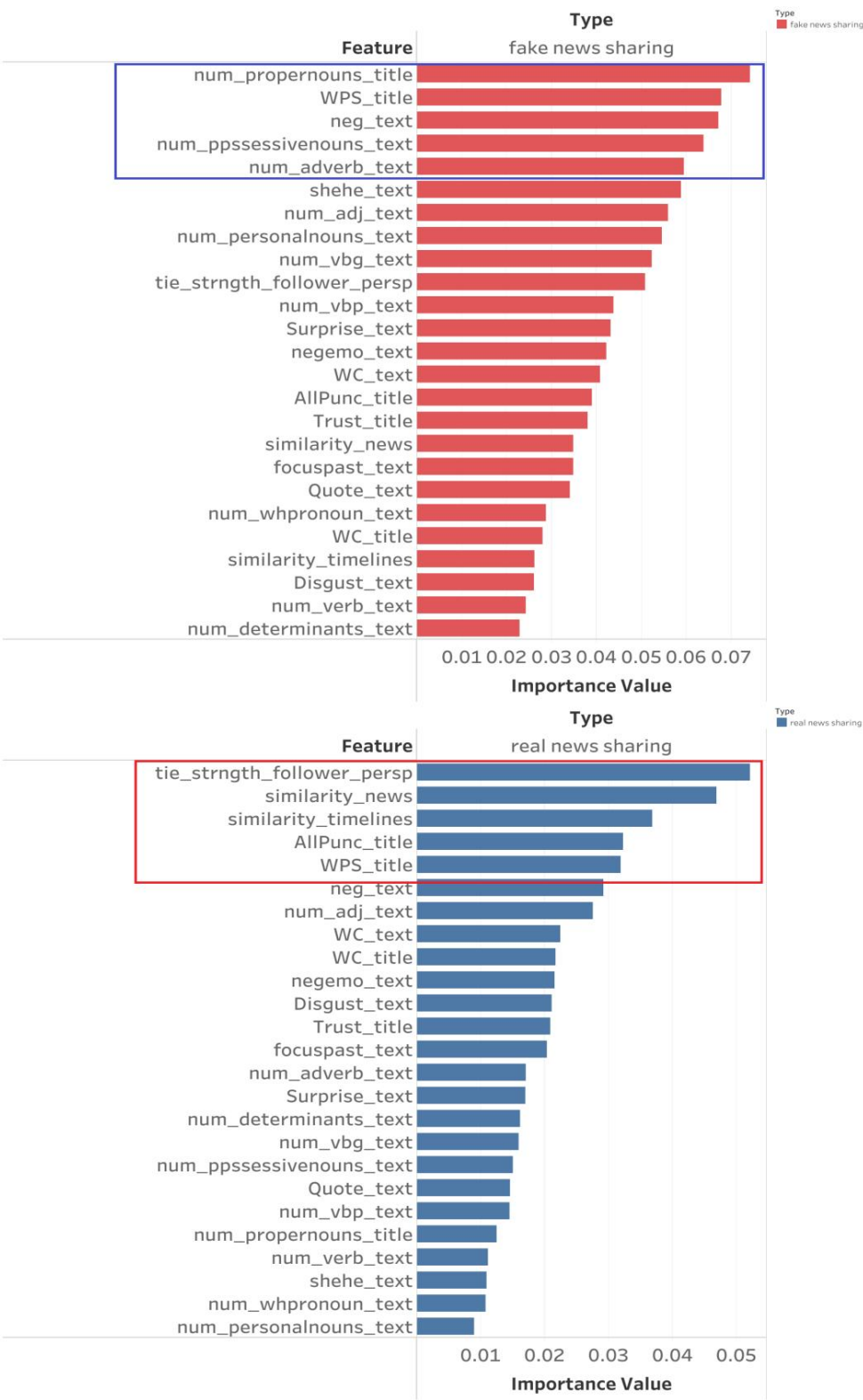
Hình 5.4: Kết quả loại bỏ tính năng (Người dùng)



Hình 5.5: Kết quả loại bỏ tính năng (Người dùng)



Hình 5.6: Mức độ quan trọng của các đặc điểm tính cách đối với việc chia sẻ tin tức giả (màu đỏ) và tin tức thật (màu xanh).



Hình 5.7: Mức độ quan trọng của các tính năng chính đối với việc chia sẻ tin tức giả (màu đỏ) và tin tức thật (màu xanh).



Hình 5.8: Kết quả loại bỏ tính năng chính được đơn giản hóa

CHƯƠNG 6:

KẾT LUẬN VÀ CÔNG VIỆC TƯƠNG LAI

Việc phát hiện tin tức giả đã trở nên quan trọng trong thời đại internet ngày nay. Nghiên cứu này thảo luận việc thực hiện một phương pháp dựa trên lý thuyết DOI để dự đoán, nếu và khi người dùng sẽ chia sẻ tin tức thật hoặc giả. Hành vi của người dùng và các yếu tố tương quan với xu hướng chia sẻ tin tức giả hay thật đã được xác định. Chia sẻ tin tức thật và giả tốt hơn được dự đoán dựa trên đặc điểm của người dùng, tin tức và mạng xã hội.

Trong nghiên cứu này, chúng tôi trình bày thiết kế và đánh giá thực tế của dự đoán của việc chia sẻ xã hội. Dựa trên lý thuyết DOI, chúng tôi đã chứng minh các yếu tố dự đoán chia sẻ tin tức giả và thật giữa người dùng. Phân tích đã hỗ trợ cho giả thuyết rằng việc chia sẻ tin tức thật và giả được dự đoán tốt hơn với người dùng, tin tức và mạng xã hội đặc trưng.

Phương pháp tiếp cận được đề xuất là kết hợp các tính năng tin tức, người dùng và mạng lưới đã được thúc đẩy AUROC tăng 30% so với các mô hình cơ sở dựa trên mạng. Trong số các tính năng chính, chúng tôi thấy rằng việc chia sẻ tin tức thật và giả cho thấy sự phụ thuộc cao vào sự tương đồng của người dùng, văn bản tin tức, tiêu đề tin tức, sức mạnh liên kết và các thuộc tính rõ ràng. Đối với phần thứ hai của nghiên cứu của chúng tôi, mô hình sống sót đã được sử dụng để dự đoán thời gian chia sẻ một tin tức nhất định của người dùng. Sai số tuyệt đối trung bình đạt được bởi ngẫu nhiên rừng sinh tồn thấp và đáng tin cậy. Mặc dù Chỉ số Concordance (C-Index) cho chia sẻ tin tức thực tế thấp hơn một chút so với mức cơ bản, Chỉ số C của Ngẫu nhiên

Survival Forest (RSF) có thể so sánh với đường cơ sở để chia sẻ tin tức giả mạo.

6.0.1 Công việc trong tương lai

Công trình này chứng minh cách dữ liệu người dùng và tin tức có thể được sử dụng để dự đoán một số khía cạnh của việc chia sẻ tin tức giả mạo. Cách tiếp cận này có thể được mở rộng hơn nữa và trở nên mạnh mẽ hơn theo nhiều cách. Ngày nay, rất nhiều thông tin được chia sẻ thông qua đồ họa như hình ảnh, gif, v.v. Phân tích các ảnh gif và hình ảnh trong tweet có thể thúc đẩy dự đoán về việc chia sẻ tin tức giả mạo và có tiềm năng nâng cao hiểu biết về tính cách. Việc đào tạo mô hình với một tập dữ liệu lớn hơn là một trong những thách thức trong công việc này. mô hình sinh tồn được sử dụng ở đây mất nhiều thời gian xử lý. Trong giai đoạn tiếp theo của công việc, một mô hình sống sót tham số có thể được khám phá để xem liệu điều này có giúp giảm thiểu thời gian đào tạo. Một khía cạnh thú vị khác có thể được khám phá là tốc độ tin tức sự lan truyền. Tin giả có xu hướng lan truyền nhanh hơn và xa hơn so với tin thật. Vì vậy, ước tính tốc độ lan truyền của tin tức có khả năng cải thiện khả năng phát hiện thông tin sai lệch trên phương tiện truyền thông xã hội. Một cơ hội khác là mở rộng nghiên cứu này các nền tảng truyền thông xã hội khác và khai thác thêm các tính năng người dùng và tin tức. Do đó, tiếp tục tối ưu hóa phương pháp luận để cải thiện thời gian xử lý, khám phá nhiều loại dữ liệu hơn, trích xuất các tham số truyền bá động của tin tức và việc mở rộng các mô hình này sang các nền tảng truyền thông xã hội khác có thể mang lại nhiều hiểu biết sâu sắc hơn dự đoán và phòng ngừa tin giả.

6.0.2 Hạn chế

Các mô hình sinh tồn được sử dụng để phân tích khoảng thời gian dự kiến cho đến khi một sự kiện xảy ra. Trong nghiên cứu này, sự kiện là quyết định của người dùng có chia sẻ hay không chia sẻ một mẫu tin tức. Một trong những nhược điểm của nghiên cứu này là gói hàng đã lấy thời gian đào tạo cao đáng kể cho tập dữ liệu chính. Vấn đề đã được giải quyết bằng cách sử dụng một tập dữ liệu nhỏ hơn, 800 trường hợp dữ liệu được lấy mẫu. Sẽ rất sáng suốt khi kiểm tra hiệu suất trên các mô hình sinh tồn khác như mô hình tham số. Các mô hình tham số mô hình có lợi thế là dễ dàng kết hợp các tính năng vào mô hình và suy luận thủ tục. Một hạn chế khác là phương pháp quy kết được sử dụng. Chúng tôi các tính năng tính toán cho nhiệm vụ phân loại và sử dụng phép quy nạp đơn giản với có nghĩa là để giải quyết các quan sát bị thiếu. Cách tiếp cận này không được khuyến khích vì nó đưa ra sự thiên vị.

Nghiên cứu cũng đã triển khai lý thuyết khuếch tán đổi mới (DOI) để tích hợp nhiều cấp độ ảnh hưởng theo người dùng, đổi mới và mạng lưới. Bộ dữ liệu của chúng tôi là nhỏ gọn, lý thuyết cần được kiểm tra với một tập dữ liệu lớn hơn với xung đột dữ liệu. Cách tiếp cận này có thể góp phần điều chỉnh lý thuyết DOI hoặc hiểu lý thuyết này có tác động đến các tập dữ liệu lớn hơn.

TÀI LIỆU THAM KHẢO

- [1] A Wolfman-Arent. Cuộc chiến đang diễn ra giữa giáo viên khoa học và tin tức giả. Đài phát thanh công cộng quốc gia (NPR): Phiên bản buổi sáng, 2017.
- [2] Lance E Mason, Dan Krutka và Jeremy Stoddard. Hiểu biết về phương tiện truyền thông, dân chủ, và thách thức của tin giả. Tạp chí Giáo dục về Kỹ năng Truyền thông, 10(2):1-10, 2018.
- [3] Ying Lin. 10 số liệu thống kê về Twitter mà mọi nhà tiếp thị nên biết vào năm 2020 [đồ họa thông tin].
- [4] Paulo Shakarian, Abhivav Bhatnagar, Ashkan Aleali, Elham Shaabani, Ruocheng Guo, et al. Sự lan truyền trong mạng xã hội. Springer, 2015.
- [5] Anna Sophie K"umpel, Veronika Karnowski và Till Keyling. Chia sẻ tin tức tại phương tiện truyền thông xã hội: Đánh giá nghiên cứu hiện tại về người dùng chia sẻ tin tức, nội dung và mạng lưới. Truyền thông xã hội + xã hội, 1(2):2056305115610141, 2015.
- [6] Everett M Rogers. Sự lan truyền của những đổi mới. Simon và Schuster, 2010.
- [7] Soroush Vosoughi, Deb Roy, và Sinan Aral. Sự lan truyền của tin tức đúng và sai trực tuyến. Khoa học, 359(6380):1146-1151, 2018.
- [8] Chei Sian Lee và Long Ma. Chia sẻ tin tức trên phương tiện truyền thông xã hội: Hiệu ứng của lòng biết ơn các hiệu ứng và kinh nghiệm trước đó. Máy tính trong hành vi của con người, 28(2):331-339, 2012.

- [9] Kai Shu, Suhang Wang, và Huan Liu. Ngoài nội dung tin tức: Vai trò của xã hội
bối cảnh để phát hiện tin tức giả mạo. Trong Biên bản của Hội nghị quốc tế ACM lần thứ mười hai
Hội nghị về Tìm kiếm trên Web và Khai thác dữ liệu, trang 312-320, 2019.
- [10] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang và Huân Liu. Tin tức giả
tection trên phương tiện truyền thông xã hội: Một góc nhìn khai thác dữ liệu. ACM SIGKDD Explorations
Bản tin, 19(1):22-36, 2017.
- [11] Kai Shu, Suhang Wang, và Huan Liu. Hiểu hồ sơ người dùng trên mạng xã hội
phương tiện truyền thông để phát hiện tin tức giả mạo. Trong Hội nghị IEEE năm 2018 về Thông tin đa phương tiện
Xử lý và Truy xuất thông tin (MIPR), trang 430-435. IEEE, 2018.
- [12] Gordon Pennycook, Ziv Epstein, Mohsen Mosleh, Antonio A Arechar, Dean
Eckles và David G Rand. Hiểu và giảm thiểu sự lan truyền thông tin sai lệch
thông tin trực tuyến, 2019.
- [13] Long Ma, Chei Sian Lee, và Dion H Goh. Hiểu về việc chia sẻ tin tức trong mạng xã hội
phương tiện truyền thông từ góc độ phổ biến của các sáng kiến. Năm 2013 IEEE International
Hội nghị về Điện toán và Truyền thông Xanh và IEEE Internet of
Things và IEEE Cyber, Physical and Social Computing, trang 1013-1020. IEEE,
2013.
- [14] Andrew Guess, Jonathan Nagler và Joshua Tucker. Ít hơn bạn nghĩ: Preva-
và các yếu tố dự báo về việc phát tán tin tức giả trên facebook. Khoa học tiến bộ,
5(1):eaau4586, 2019.
- [15] Waheeb Yaqub, Otari Kakhidze, Morgan L Brockman, Nasir Memon, và
Sameer Patil. Tác động của các chỉ số uy tín đến ý định chia sẻ tin tức trên mạng xã hội.

Trong Biên bản báo cáo Hội nghị CHI năm 2020 về Yếu tố con người trong Điện toán Hệ thống, trang 1-14, 2020.

- [16] Kai Shu, H Russell Bernard, và Huan Liu. Nghiên cứu tin tức giả mạo qua mạng phân tích: phát hiện và giảm thiểu. Trong Những thách thức nghiên cứu mới nổi và Op-cơ hội trong Phân tích và Khai thác Mạng xã hội Tính toán, trang 43-65.

Nhà xuất bản Springer, 2019.

- [17] Eben Kenah. Phân tích sự sống còn phi tham số của dữ liệu bệnh truyền nhiễm. Tạp chí của Hiệp hội Thống kê Hoàng gia: Loạt B (Phương pháp Thống kê), 75(2):277-303, 2013.

- [18] David Kempe, Jon Kleinberg và Eva Tardos. Tối đa hóa sự lan truyền của sự ảnh hưởng thông qua mạng xã hội. Trong Biên bản của ACM SIGKDD lần thứ chín hội nghị quốc tế về Khám phá tri thức và khai thác dữ liệu, trang 137-146, 2003.

- [19] Kai Shu, Deepak Mahudeswaran và Huan Liu. Fakenewstracker: một công cụ để làm giả thu thập tin tức, phát hiện và trực quan hóa. Tính toán và Toán học Lý thuyết tổ chức, 25(1):60-71, 2019.

- [20] Yair Neuman. Phân tích tính cách tính toán: Giới thiệu, ứng dụng thực tế cation và hướng đi mới. Springer, 2016.

- [21] Libby Hemphill, Aron Culotta và Matthew Heston. Điểm cực: Đo lường đảng phái sử dụng nội dung phương tiện truyền thông xã hội. Tạp chí Công nghệ thông tin & Chính trị, 13(4):365-377, 2016.

- [22] Joshua M Chamberlain, Francesca Spezzano, Jaclyn J Kettler và Bogdan Dit. Phân tích mạng lưới tương tác trên Twitter của các thành viên quốc hội Hoa Kỳ. ACM Giao dịch về máy tính xã hội, 4(1):1-22, 2021.
- [23] Zijian Wang, Scott Hale, David Ifeoluwa Adelani, Przemyslaw Grabowicz, Timo Hartman, Fabian Flöck và David Jurgens. Suy luận và phản hồi nhân khẩu học ước tính dân số phản hồi từ dữ liệu phương tiện truyền thông xã hội đa ngôn ngữ. Trong Hội nghị World Wide Web, trang 2056-2067, 2019.
- [24] Sharath Chandra Guntuku, Anneke Buffone, Kokil Jaidka, Johannes C Eichstaedt, và Lyle H Ungar. Hiểu và đo lường căng thẳng tâm lý sử dụng phương tiện truyền thông xã hội. Trong Biên bản Hội nghị AAAI quốc tế về Web và phương tiện truyền thông xã hội, tập 13, trang 214-225, 2019.
- [25] James W Pennebaker, Martha E Francis và Roger J Booth. Ngôn ngữ học câu hỏi và số lượng từ: Liwc 2001. Mahway: Lawrence Erlbaum Associates, 71(2001):2001, 2001.
- [26] Yla R Tausczik và James W Pennebaker. Ý nghĩa tâm lý của từ ngữ: Liwc và phương pháp phân tích văn bản vi tính. Tạp chí ngôn ngữ và xã hội tâm lý học, 29(1):24-54, 2010.
- [27] Wei Wang, Ivan Hernandez, Daniel A Newman, Jibo He và Jiang Bian. Twitter: Phân tích ter: Nghiên cứu xu hướng hàng tuần của chúng tôi về căng thẳng công việc và cảm xúc. Ứng dụng Tâm lý học, 65(2):355-378, 2016.
- [28] Radim Rehůřek và Petr Sojka. Khung phần mềm cho mô hình chủ đề với Tập đoàn lớn. Trong Biên bản của Hội thảo LREC 2010 về Những thách thức mới

cho NLP Frameworks, trang 45-50, Valletta, Malta, tháng 5 năm 2010. ELRA. <http://is.muni.cz/publication/884893/en>.

[//is.muni.cz/publication/884893/en](http://is.muni.cz/publication/884893/en).

[29] Saif M Mohammad. Cường độ ảnh hưởng của từ ngữ. bản in trước arXiv arXiv:1704.08798, 2017.

[30] Ashlee Milton, Leveson Batista, Garrett Allen, Siqi Gao, Yiu-Kai D Ng, và Maria Soledad Pera. “đừng đánh giá một cuốn sách qua trang bìa”: Khám phá những đặc điểm của cuốn sách trẻ em ủng hộ. Trong Hội nghị ACM lần thứ mười bốn về Hệ thống khuyến nghị, các trang 669-674, 2020.

[31] Benjamin Horne và Sibel Adali. Tin mới nhất: Tin giả chứa nhiều thứ trong tiêu đề, sử dụng nội dung đơn giản, lặp lại trong nội dung văn bản, giống với châm biếm hơn là thực tế tin tức. Trong Biên bản Hội nghị quốc tế AAAI về Web và Xã hội Media, tập 11, năm 2017.

[32] Anu Shrestha và Francesca Spezzano. Đặc điểm văn bản của tiêu đề tin tức và cơ thể để phát hiện tin tức giả: Một nghiên cứu khả năng tái tạo. Trong Djoerd Hiemstra, Marie-Francine Moens, Josiane Mothe, Raffaele Perego, Martin Potthast và Fabrizio Sebastiani, biên tập viên, Những tiến bộ trong việc truy xuất thông tin - Hội nghị châu Âu lần thứ 43 Hội nghị về Nghiên cứu IR, ECIR 2021, Sự kiện trực tuyến, 28 tháng 3 - 1 tháng 4, 2021, Biên bản, Phần II, tập 12657 của Ghi chú bài giảng về Khoa học máy tính, trang 120-133. Springer, 2021.

[33] Bilal Ghanem, Paolo Rosso và Francisco Rangel. Một phân tích cảm xúc về sự sai lầm thông tin trên phương tiện truyền thông xã hội và các bài báo. Giao dịch ACM trên Internet Công nghệ (TOIT), 20(2):1-18, 2020.

- [34] G Harry Mc Laughlin. Phân loại khói bụi - công thức để đọc mới. Tạp chí
đọc, 12(8):639-646, 1969.
- [35] Olaf Zorzi. Granovetter (1983): Sức mạnh của các mối quan hệ yếu: Một lý thuyết mạng lưới
xem lại. Trong Schlüsselwerke der Netzwerkforschung, trang 243-246. mùa xuân,
2019.
- [36] Hemant Ishwaran, Udaya B Kogalur, Eugene H Blackstone, Michael S Lauer,
et al. Rừng sống sót ngẫu nhiên. Biên niên sử thống kê ứng dụng, 2(3):841-860,
2008.
- [37] Ping Wang, Yan Li, và Chandan K Reddy. Học máy để phân tích sự sống còn
ysis: Một cuộc khảo sát. Khảo sát máy tính ACM (CSUR), 51(6):1-36, 2019.
- [38] Amit Goyal, Francesco Bonchi và Laks VS Lakshmanan. Ảnh hưởng của việc học
khả năng trong mạng xã hội. Trong Biên bản báo cáo của ACM quốc tế lần thứ ba
hội nghị về tìm kiếm trên Web và khai thác dữ liệu, trang 241-250, 2010.
- [39] Yuchen Li, Ju Fan, Yanhao Wang và Kian-Lee Tan. Tối đa hóa ảnh hưởng đến
biểu đồ xã hội: Một cuộc khảo sát. Giao dịch IEEE về Kỹ thuật dữ liệu và kiến thức,
30(10):1852-1872, 2018.
- [40] Stephane Fotso et al. PySurvival: Gói nguồn mở để phân tích sự sống còn
mô hình hóa, 2019-.