**TRƯỜNG ĐẠI HỌC SÀI GÒN**



**KHOA CÔNG NGHỆ THÔNG TIN**



**TIỀU LUẬN: PHƯƠNG PHÁP NGHIÊN CỨU KHOA HỌC TRONG CNTT**

**TÊN ĐỀ TÀI: NGHIÊN CỨU ỨNG DỤNG KỸ THUẬT LỌC DỰA TRÊN NỘI DUNG VÀO HỆ THỐNG GỢI Ý SẢN PHẨM TRÊN SÀN THƯƠNG MẠI ĐIỆN TỬ LAZADA**

**Trần Phan Thanh Long**

**GIÁO VIÊN HƯỚNG DẪN: TS. ĐINH THỊ THU HƯƠNG**

**TP. HCM tháng 11/2020**

**TRƯỜNG ĐẠI HỌC SÀI GÒN**

**KHOA CÔNG NGHỆ THÔNG TIN**

**ĐỀ CƯƠNG KHÓA LUẬN TỐT NGHIỆP**

**I – Phần thông tin chung**

**Họ và tên sinh viên :** Trần Phan Thanh Long

MSSV : 3118410248 Lớp:DTC11811

E-mail : tranphanthanhlong18@gmail.com

Ngành : Công nghệ thông tin

Chuyên ngành : Công nghệ phần mềm

**Tên đề tài**: Nghiên cứu ứng dụng kỹ thuật lọc dựa trên nội dung vào hệ thống gợi ý sản phẩm trên sàn thương mại điện tử Lazada

**Giảng viên hướng dẫn**: Giảng Viên Đinh Thị Thu Hương.

**II - Nội dung đề cương**

(Mẫu đề cương tham khảo, GVHD có thể thay đổi cho phù hợp với đề tài)

1. Tính cấp thiết của đề tài/ Lý do chọn đề tài:

Hệ thống gợi ý (recommender system) được coi là ứng dụng phổ biến nhất của lĩnh vực big data giúp người dùng tìm kiếm thông tin có liên quan trong thời đại dữ liệu tràn lan. Đó là kết quả của sự mở rộng trong lĩnh vực kinh doanh thương mại điện tử, khách hàng (người mua) phải xử lý một lượng lớn thông tin trước khi họ quyết định mua một mặt hàng. Hệ thống gợi ý cung cấp giải pháp cho vấn đề quá tải thông tin này vì nó được sử dụng để gợi ý sản phẩm cho khách hàng. Nó nâng cao hoạt động kinh doanh thương mại điện tử bằng việc tìm ra những gợi ý thú vị cho người dùng khi họ có ý định mua nhưng không thể xem qua hết tất cả giang hàng.

1. Mục đích KLTN:

Cải thiện trải nghiệm mua sắm của người tiêu dùng trên nền tảng thương mại điện tử. Từ đó góp phần mở rộng sức ảnh hưởng của thương mại điện tử.

1. Phạm vi nghiên cứu: sàn thương mại điện tử lazada
2. Phương pháp nghiên cứu:

Các phương pháp xử lý thông tin:

-Phương pháp nghiên cứu lý thuyết

-Phương pháp thực nghiệm

-Phương pháp xử lý thông tin

1. Cấu trúc của KLTN:

**Chương 1: Tổng quan**

1.1. Tầm quan trọng và ý nghĩa của hệ thống gợi ý

1.2. Tổng quan vấn đề của hệ thống gợi ý còn gặp phải

1.3. Mục tiêu đề tài

1.4. Đối tượng và phạm vi đề tài

1.5. Phương pháp nghiên cứu, thực hiện đề tài

**Chương 2: Lý thuyết cơ sở của giải thuật lọc dựa trên nội dụng**

2.1. Các giải thuật gợi ý phổ biến

2.2. Cơ sở của giải thuật lọc theo nội dung

**Chương 3: Thiết kế giải thuật lọc dựa trên nội dung**

**Chương 4: Thử nghiệm và đánh giá kết quả**

4.1. Tập dữ liệu được dùng cho thử nghiệm thuật toán

4.2. Triển khai thuật toán trên tập dữ liệu

4.3. Kết quả chạy thử nghiệm

4.4. Đánh giá kết quả

**Chương 5: Kết luận và hướng phát triển**

5.1. Kết luận

5.2 Hướng phát triển

1. Kế hoạch thực hiện đồ án/ khóa luận tốt nghiệp trong 10 tuần:

|  |  |  |
| --- | --- | --- |
| **Tuần** | **Công việc dự kiến** | **Ghi chú** |
| 1 | Khảo sát những vấn đề gặp phải trong hệ thống gợi ý ở nền tản thương mại điện tử lazada gặp phải |  |
| 2 | Nghiên cứu tài liệu về hệ thống gợi ý dựa trên nội dung |  |
| 3 | Phát triển thuật toán |  |
| 4 | Triển khi thuật toán trên phần cứng |  |
| 5 | Chuẩn bị tập dữ liệu |  |
| 6 | Triển khai chạy thuật toán trên tập dữ liệu và ghi nhận kết quả |  |

*TP. HCM, ngày … tháng … năm …*

**Ý kiến giảng viên hướng dẫn Sinh viên thực hiện**

*(Ký và ghi rõ họ tên) (Ký và ghi rõ họ tên)*

**Ý kiến của trưởng Bộ môn**

*(Ký và ghi rõ họ tên)*

|  |  |
| --- | --- |
|  |  |

Phiếu khảo sát mức độ hữu dụng của hệ thống gợi ý lazada

Chào anh/chị. Em xin chút thời gian của anh/chị dành ra để chia sẻ trải nghiệm của bản thân trong trong việc mua hàng trên lazada nha!!!

1. Anh/chị thường lướt XEM sản phẩm trên lazada vào các dịp nào?

*Chọn tất cả mục phù hợp.*

Khi cần mua sản phẩm đó

Khi có thông báo khuyến mãi cho cá nhân

Trong các sự kiện sale lớn Khi rảnh

Không bao gi

ờ

Th

ư

ờ

ng xuyên

1. Anh/chị thường QUYẾT ĐỊNH mua sản phẩm trên lazada vào các dịp nào?

*Chọn tất cả mục phù hợp.*

Khi cần mua sản phẩm đó

Khi có thông báo khuyến mãi cho cá nhân

Trong các sự kiện sale lớn

Khi rảnh

1. Anh/chị có thường xuyên vào xem sản phẩm được lazada gợi ý trên facebook không?

*Chỉ đánh dấu một hình ôvan.*

* 1. 2 3 4 5

1. Anh/chị có thường xuyên quyết định mua luôn sản phẩm được lazada gợi ý trên facebook không?

*Chỉ đánh dấu một hình ôvan.*

* 1. 2 3 4 5

Không bao gi

ờ

Th

ư

ờ

ng xuyên

1. Khi đang xem một sản phẩm anh chị có thường xuyên kéo xuống dưới để xem danh sách sản phẩm được gợi ý không?

*Chỉ đánh dấu một hình ôvan.*

* 1. 2 3 4 5

Không bao gi

ờ

Th

ư

ờ

ng xuyên

1. Anh/chị có thường xuyên mua được sản phẩm ưng ý thông qua danh sách gợi ý của lazada không?

*Chỉ đánh dấu một hình ôvan.*

* 1. 2 3 4 5

Không bao gi

ờ

Th

ư

ờ

ng xuyên

1. Điều gì làm anh chị quyết định không tiếp tục lướt xem và mua hàng trên lazada? Do tìm không ra sản phẩm mình cần mua? Hay do tìm hết tất cả các shop bán sản phẩm đó nhưng giá thì trên trời? Hay vì bất kỳ lý do nào khác ạ?

Cảm ơn anh chị đã bỏ thời gian quý báo để hoàn thành phiếu khảo sát!!! Chúc anh chị có được trải nghiệm mua hàng tốt nhất.

XỬ LÝ KẾT QUẢ ĐIỀU TRA

- Sự tương quan giữ mức độ thường xuyên sử dụng hệ thống gợi ý và mức độ mua hàng thông qua hệ thống gợi ý

Bảng thống kê

|  |  |  |
| --- | --- | --- |
| **STT** | **Mức độ thường xuyên sử dụng hệ thống gợi ý** | **Mức độ mua hàng thông qua hệ thống gợi ý** |
|  | 2 | 2 |
|  | 4 | 3 |
|  | 5 | 4 |
|  | 2 | 5 |
|  | 5 | 4 |
|  | 2 | 3 |
|  | 2 | 2 |
|  | 5 | 3 |
|  | 5 | 4 |
|  | 5 | 2 |
|  | 3 | 5 |
|  | 3 | 2 |
|  | 2 | 4 |
|  | 4 | 5 |
|  | 4 | 2 |
|  | 4 | 2 |
|  | 4 | 4 |
|  | 5 | 3 |
|  | 3 | 4 |
|  | 5 | 4 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **STT** | **Mức độ thường xuyên sử dụng hệ thống gợi ý** | **Thứ bậc**  **(X)** | **Mức độ mua hàng thông qua hệ thống gợi ý** | **Thứ bậc**  **(Y)** | **(X-Y)2** |
|  | 2 | 18.5 | 2 | 17.5 | **0.25** |
|  | 4 | 10 | 3 | 12.5 | **6.25** |
|  | 5 | 3.5 | 4 | 7 | **9** |
|  | 4 | 10 | 5 | 2 | **256** |
|  | 5 | 3.5 | 4 | 7 | **9** |
|  | 2 | 18.5 | 3 | 12.5 | **30.25** |
|  | 2 | 18.5 | 2 | 17.5 | **0.25** |
|  | 4 | 10 | 3 | 12.5 | **72.25** |
|  | 5 | 3.5 | 4 | 7 | **9** |
|  | 3 | 15 | 2 | 17.5 | **182.25** |
|  | 4 | 10 | 5 | 2 | **144** |
|  | 3 | 15 | 2 | 17.5 | **12.25** |
|  | 5 | 3.5 | 4 | 7 | **121** |
|  | 4 | 10 | 5 | 2 | **64** |
|  | 2 | 18.5 | 2 | 17.5 | **56.25** |
|  | 4 | 10 | 2 | 17.5 | **56.25** |
|  | 4 | 10 | 4 | 7 | **9** |
|  | 5 | 3.5 | 3 | 12.5 | **72.25** |
|  | 3 | 15 | 4 | 7 | **49** |
|  | 5 | 3.5 | 4 | 7 | **0.25** |
|  |  |  |  | **Tổng** | **1158.5** |

**Hệ số tương quan thứ bậc (giữa mức đọ sử dụng và mức độ mua)**

R =

**=** 0.1289473684210526

**Kết luận:** mức độ thường xuyên sử dụng hệ thống gợi ý và mức độ mua hàng dựa trên hệ thống gợi ý tương quan thuận với nhau. Nghĩa là hệ thống gợi ý hiện tại của lazada cần được cải thiện để tăng giá trị hệ số tương quan. Hệ số tương quan tỉ lệ thuận với mức độ hiệu quả của hệ thống gợi ý.

- Sự tương quan giữ mức độ thường xuyên vào xem sản phẩm được gợi ý trên facebook và mức độ sử dụng hệ thống gợi ý

Bảng thống kê

|  |  |  |
| --- | --- | --- |
| **STT** | **Mức độ thường xuyên vào xem sản phẩm được gợi ý trên facebook** | **Mức độ thường xuyên sử dụng hệ thống gợi ý** |
|  | 2 | 2 |
|  | 4 | 4 |
|  | 5 | 5 |
|  | 4 | 2 |
|  | 5 | 5 |
|  | 5 | 2 |
|  | 4 | 2 |
|  | 3 | 5 |
|  | 3 | 5 |
|  | 3 | 5 |
|  | 3 | 3 |
|  | 4 | 3 |
|  | 4 | 2 |
|  | 5 | 4 |
|  | 4 | 4 |
|  | 5 | 4 |
|  | 3 | 4 |
|  | 4 | 5 |
|  | 5 | 3 |
|  | 4 | 5 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **STT** | **Mức độ thường xuyên vào xem sản phẩm được gợi ý trên facebook** | **Thứ bậc**  **(X)** | **Mức độ thường xuyên sử dụng hệ thống gợi ý** | **Thứ bậc**  **(Y)** | **(X-Y)2** |
|  | 2 | 20 | 2 | 18 | **0.25** |
|  | 4 | 10.5 | 4 | 10 | **0** |
|  | 5 | 3.5 | 5 | 4 | **0.25** |
|  | 4 | 10.5 | 2 | 18 | **64** |
|  | 5 | 3.5 | 5 | 4 | **0.25** |
|  | 5 | 3.5 | 2 | 18 | **0.25** |
|  | 4 | 10.5 | 2 | 18 | **0.25** |
|  | 3 | 17 | 5 | 4 | **36** |
|  | 3 | 17 | 5 | 4 | **0.25** |
|  | 3 | 17 | 5 | 4 | **121** |
|  | 3 | 17 | 3 | 14 | **16** |
|  | 4 | 10.5 | 3 | 14 | **1** |
|  | 4 | 10.5 | 2 | 18 | **210.25** |
|  | 5 | 3.5 | 4 | 10 | **0** |
|  | 4 | 10.5 | 4 | 10 | **72.25** |
|  | 5 | 3.5 | 4 | 10 | **0** |
|  | 3 | 17 | 4 | 10 | **0** |
|  | 4 | 10.5 | 5 | 4 | **0.25** |
|  | 5 | 3.5 | 3 | 14 | **1** |
|  | 4 | 10.5 | 5 | 4 | **0.25** |
|  |  |  |  | **Tổng** | **523.5** |

**Hệ số tương quan thứ bậc**

R =

**=** 0.60639097744

**Kết luận:** mức độ thường xuyên xem sản phẩm gợi ý trên facebook và mức độ sử dụng hệ thống gợi ý tương quan với nhau. Điều này thể hiện việc người dùng facebook quan tâm đến sản phẩm được lazada gợi ý phần lớn đều sử dụng hệ thống gợi ý của lazada.

# Bibliography

|  |  |  |  |
| --- | --- | --- | --- |
| [1] |  |  | B. Saumya, "Researchgate," International Journal of Knowledge and Systems Science, April 2019. [Online]. Available: https://www.researchgate.net/publication/334577400\_A\_Study\_of\_Recent\_Recommender\_System\_Techniques. |
| [2] |  |  | D. T. T. Huong, "PPNCKH nhóm 3| thứ 5| tiết 1-2," Facebook, 3 September 2020. [Online]. Available: https://www.facebook.com/groups/596954840998030/files. |