

# XÂY DỰNG HỆ THỐNG GỢI Ý CỦA HÀNG

THỰC TẬP DỰ ÁN TỐT NGHIỆP

GIÁO VIÊN HƯỚNG DẪN ThS Tiết Gia Hồng

Nguyễn Quốc Anh 20120429 Tô Trần Sơn Bá 20120431 Hoàng Văn Cầu 20120439 Nguyễn Đình Cường 20120446 Nguyễn Trung Hiếu 20120477 Vương Tân Phát 20120344

01 - GIỚI THIỆU
05 - KIẾN TRÚC HỆ THỐNG
06 - SẢN PHẨM

**498,8 ngàn tỷ đồng**  
doanh thu năm 2023

**61 triệu**  
người dùng mua sắm

**7,5 triệu đồng/năm**  
mức mua trung bình một người

Thời gian chọn cửa hàng để mua sản phẩm

Tỷ lệ phản hồi người dùng khi nhận sản phẩm không giống mô tả

**Cửa hàng**

- Tổng số lượt theo dõi
- Rating trung bình của cửa hàng
- Số lượng sản phẩm đã bán
- Cửa hàng chính hãng

**Người dùng:**

- Rating đánh giá sản phẩm
- Rating của bình luận
- Hành vi người dùng

$\text{cosine\_similarity}(a, b) = \frac{a \cdot b}{|a| \cdot |b|}$

$$\hat{R}_{u_i, s_j} = \mu_{u_j} + \frac{\sum_{p \in P_{u_i, s_j}} \text{sim}_{p, u_i} (R_{p, s_j} - \mu_p)}{\sum_{p \in P_{u_i, s_j}} \text{sim}_{p, u_i}}$$

02 - KHẢO SÁT ỨNG DỤNG
03 - MỤC TIÊU
04 - QUÁ TRÌNH THỰC HIỆN

**TIKI.VN**

**Lazada**

**Shopee**

3/3 hỗ trợ danh sách sản phẩm gợi ý

3/3 không hỗ trợ tích hợp tư vấn vào tìm kiếm

**TÍNH NĂNG CHÍNH : Gợi ý cửa hàng**

**TÍNH NĂNG PHỤ : Gợi ý sản phẩm liên quan đến những sản phẩm người dùng xem gần đây**

05 - KIẾN TRÚC HỆ THỐNG
06 - SẢN PHẨM

**Phân hệ Admin**

**Phân hệ Client**

**Hệ thống thương mại điện tử ShopLY:** [shoply-7ixf.onrender.com](http://shoply-7ixf.onrender.com)

07 - KẾT QUẢ ĐẠT ĐƯỢC
08 - HẠN CHẾ

**THỬ NGHIỆM ELASTICSEARCH**

1. 1 người dùng gửi 100 yêu cầu tìm kiếm sản phẩm với từ khóa tìm kiếm là 'áo'

100% thời gian phản hồi dưới 1s

Trung bình thời gian phản hồi 0,2s

100 phản hồi hoàn thành trong 24s

**Thử nghiệm**

7784 yêu cầu tìm kiếm được thực hiện

Trung bình yêu cầu được xử lý trong 1s

Tỉ lệ thất bại 0%

Thời gian phản hồi nhanh nhất 0,5s

Thời gian phản hồi chậm nhất 6,6s

09 - HƯỚNG PHÁT TRIỂN

**DỮ LIỆU**

Tiến hành crawl dữ liệu từ **TIKI.VN**

**QUÁ TRÌNH THỰC HIỆN**

- 01 KHẢO SÁT
- 02 DỮ LIỆU
- 03 TÌM HIỂU
- 04 PHÁT HÀNH
- 05 CÀI ĐẶT

Thuật toán		
NEIGHBORHOOD-BASED CONTENT-BASED MATRIX-FACTORIZATION		
Mô hình ngôn ngữ	Dữ liệu	Elasticsearch
PhoBERT	31,432	
Tổng	1,936,395	

**Áp dụng các thuật toán tư vấn vào gợi ý cửa hàng**

**Dữ liệu sử dụng trong thuật toán gợi ý cửa hàng chưa đa dạng**

Ma trận Rating: 2034 người dùng x 9026 cửa hàng = 18,358,884 rating

Dữ liệu chỉ có 9414 rating

Mô hình PhoBERT chưa đánh giá chính xác hoàn toàn các câu bình luận do có thể có nhiều khía cạnh trong câu

Chưa giải quyết được khuyết điểm của Neighborhood-based như người dùng mới, sản phẩm mới, ...

**Sử dụng dữ liệu từ mạng xã hội để gợi ý sâu hơn cho người dùng**

**Sử dụng phương pháp lai giữa Neighborhood-based với Knowledge-based để giảm khuyết điểm của Neighborhood-based**

**Chuyển đổi cách thức tổ chức dữ liệu sang sử dụng GraphQL để tăng tốc độ tìm kiếm**

**Sử dụng mô hình ngôn ngữ lớn để đánh giá được khía cạnh của câu bình luận**