TÓM TẮT

Hệ thống tư vấn đóng vai trò quan trọng trong việc hỗ trợ người dùng lựa chọn sản phẩm phù hợp, góp phần nâng cao trải nghiệm cá nhân trên các nền tảng thương mại điện tử. Tuy nhiên, các phương pháp hiện tại gặp nhiều thách thức do dữ liệu rating thường thiếu chi tiết, không đầy đủ và đôi khi không phản ánh chính xác sở thích thực sự của người dùng. Để khắc phục vấn đề này, đề tài đề xuất sử dụng các bình luận của người dùng nhằm cải thiện độ chính xác trong việc ước lượng độ tương đồng sở thích. Cụ thể, nhóm tác giả áp dụng mô hình BERT để học các vector biểu diễn người dùng và sản phẩm từ nội dung bình luận. Những vector này sau đó được kết hợp với dữ liệu rating để xây dựng một hàm mục tiêu giúp đánh giá mức độ tương đồng giữa các người dùng một cách toàn diện hơn, thay vì chỉ dựa vào rating như các phương pháp truyền thống.

Bên cạnh việc nâng cao độ chính xác, đề tài còn tập trung tối ưu hóa khả năng mở rộng của hệ thống. Nhóm tác giả đề xuất phương pháp gom cụm người dùng, giúp giảm chi phí tính toán bằng cách nhóm các người dùng có đặc điểm tương đồng lại với nhau trước khi thực hiện quá trình tư vấn. Đồng thời, nghiên cứu cũng triển khai các thuật toán tính toán theo hướng song song hóa trên nền tảng Hadoop, giúp tăng tốc độ xử lý và đảm bảo hệ thống hoạt động hiệu quả ngay cả với dữ liệu lớn.