|  |
| --- |
| Slide 1: Start (15s) |
| Thưa quý thầy cô trong hội đồng cùng toàn thể các bạn!  Hôm nay, nhóm em xin trình bày đề tài khóa luận “Ứng dụng hỗ trợ khách du lịch trên Anroid”  Đề tài này được thầy Huỳnh Hữu Việt hướng dẫn và nhóm thực hiện gồm 2 SV: Minh Hiếu, Trọng Hiếu. |
| Slide 2: Nội dung trình bày (20s) |
| Nội dung trình bày gồm 5 phần:   1. Giới thiệu đề tài 2. Khảo sát các PP khuyến nghị 3. Phương pháp thu giảm số chiều 4. Hiện thực hóa 5. Đánh giá và kết luận |
| Slide 3: Vào phần Giới thiệu đề tài (2s) |
| Sau đây là phần giới thiệu đề tài |
| Slide 4: Giới thiệu đề tài (45s) |
| * Nước ta có tiềm năng du lịch lớn, thu hút khách du lịch từ nhiều nơi trên thế giới. * ngành du lịch những năm qua ngày càng phát triển mạnh đem lại các lợi ích to lớn về kinh tế - xã hội, đóng góp lớn vào GDP của đất nước. * Ngày nay CNTT phát triển mạnh, nhiều ứng dụng và dịch vụ ra đời nhằm quảng bá hình ảnh du lịch của đất nước. Thêm nữa, xu hướng điện thoại thông minh ngày càng được sử dụng phổ biến. Với lợi thế nhỏ gọn, tiện dụng và thông minh, điện thoại thông minh hỗ trợ rất nhiều cho con người trong các hoạt động hằng ngày. * Khách du lịch cần những thông tin, những lời gợi ý chính xác, nhanh chóng tùy theo điều kiện ngữ cảnh. * Vì vậy nhóm quyết định chọn đề tài *xây dựng hệ thống khuyến nghị theo ngữ cảnh để gợi ý các địa điểm du lịch cho du khách dùng điện thoại Android.* * *(pham vi địa lý là thành phố HCM)* |
| Slide 5: Quá trình thực hiện (30s) |
| Trên màn hình là các bước chính để thực hiện đề tài:   1. Khảo sát các PP khuyến nghị để lựa chọn hướng đi thích hợp. 2. Khảo sát các kỹ thuật để triển khai hệ thống: cơ sở dữ liệu, dịch vụ web, nền tảng di động … 3. Khảo sát thông tin ngữ cảnh để biết người dùng quan tâm điều kiện ngữ cảnh nào khi đi du lịch. Cũng như thông tin các điểm du lịch nổi bật trong Tp.HCM để xây dựng cơ sở dữ liệu. 4. Hiện thực hóa hệ thống khuyến nghị và ứng dụng Android. 5. Thu thập dữ liệu đánh giá thực tế từ người dùng để phục vụ việc đánh giá thuật toán. 6. Thực nghiệm và đánh giá kết quả đạt được. |
| Slide 6: Vào phần khảo sát các PP khuyến nghị (2s) |
| Sau đây là phần khảo sát các PP khuyến nghị. |
| Slide 7: PP khuyến nghị 2 chiều (45s) |
| * Các hệ thống khuyến nghị truyền thống chỉ quan tâm 2 yếu tố (2 chiều): người dùng và đối tượng cần được khuyến nghị. Mục tiêu là dựa vào những chỉ số đánh giá đã có của người dùng đối với các đối tượng, từ đó dự đoán những chỉ số đánh giá còn thiếu. Sau đó, chọn ra những đối tượng có chỉ số đánh giá dự đoán cao nhất, gợi ý cho người dùng. * Ví dụ ta có ma trận Users x Items, có những ô người dùng đã đánh giá rồi (như người dùng u1 đánh giá đối tượng i1 giá trị là 5), có những ô chấm hỏi là chưa đánh giá, cần được dự đoán. Từ đó mới đưa ra những đối tượng có chỉ số dự đoán cao nhất và gợi ý cho người dùng. |
| Slide 8: PP khuyến nghị 2 chiều (45s) |
| **PP dựa trên nội dung:** dự đoán chỉ số đánh giá của người dùng *u* đối với đối tượng *i* dựa vào những chỉ số dự đoán của người dùng đó đối với những đối tượng tập hợp các đối tượng khác (có đặc điểm tương tự với đối tượng *i)*.  **PP đánh giá độ tương đồng*:*** *có 2 loại:*  **Dựa trên kinh nghiệm:** dự đoán chỉ số đánh giá *R(u,i)* của người dùng *u* đối với đối tượng *i* dựa vào những chỉ số dự đoán trước đây của những người dùng khác (có đặc điểm tương tự với người dùng *u*) đối với cùng đối tượng i đó.  **Dựa trên mô hình:** dùng các mô hình máy học, các thuật toán phân lớp, gom cụm …  **PP lai:** kết hợp 2 PP dựa trên nội dung và dựa trên độ tương đồng. |
| **Slide 9: Ví dụ PP 2 chiều (20s)** |
| Amazon: hệ khuyến nghị sách  Youtube: khuyến nghị các video clip  Netflix: khuyến nghị phim  Pandora: khuyến nghị các bản nhạc |
| Slide 10: PP khuyến nghị nhiều chiều (45s) |
| * Khác với PP 2 chiều (người dùng và đối tượng). PP nhiều chiều quan tâm thêm đến các chiều ngữ cảnh. * *Thông tin ngữ cảnh là những thông tin có thể mô tả được hoàn cảnh của một thực thể. VD* *Những thông tin ngữ cảnh (thời gian, thời tiết, tâm trạng, bạn đồng hành …) có ảnh hưởng đến đánh giá của người dùng.* * *Nói ví dụ blah blah blah* |
| Slide 11: Ví dụ PP khuyến nghị nhiều chiều (45s) |
| * VD không gian khuyến nghị 3 chiều (users, items, time), ta có khối lập phương chứa các chỉ số đánh giá. * Giải thích R(101,7,1) * Có những ô còn trống (do người dùng chưa đánh giá trong điều kiện thời gian đó) cần được dự đoán.   🡪 Và thuật toán khuyến nghị nhóm xây dựng là khuyến nghị trong không gian nhiều chiều sẽ được trình bày ở phần tiếp theo. |
| Slide 12: Vào phần PP thu giảm số chiều.(2s) |
| Sau đây là phương pháp khuyến nghị nhiều chiều dựa vào các thu giảm số chiều. |
| Slide 13: PP khuyến nghị nhiều chiều (45s) |
| Với PP nhiều chiều, người ta chia làm 3 loại:  (a): chọn lọc không gian dữ liệu tương ứng với điều kiện ngữ cảnh được chọn, khi đó ta được ma trận 2 chiều (Users x Items) chứa những chỉ số dự đoán ứng với điều kiện ngữ cảnh trên. Bài toán trở về 2 chiều.  (b): phớt lờ những điều kiện ngữ cảnh ban đầu, xem như không gian khuyến nghị ở đây là 2 chiều. Sau khi gợi ý các đối tượng, lọc ra những đối tượng phù hợp với điều kiện ngữ cảnh được xem xét.  (c): xây dựng những mô hình dự đoán sử dụng các điều kiện ngữ cảnh ngay lúc dự đoán chỉ số đánh giá. Chứ điều kiện ngữ cảnh ở đây không chỉ dùng ở các bước lọc như (a) và (b). |
| Slide 14: PP thu giảm số chiều (45s) |
| Ví dụ ta có khối lập phương chứa các chỉ số dự đoán của những người dùng đối với những đối tượng trong điều kiện ngữ cảnh là thời gian.  Ta thu giảm số chiều bằng cách chia khối lập phương thành 2 khối nhỏ hơn tương ứng với từng điều kiện thời gian 1 và 2. Ta gọi đó là những phân khúc dữ liệu. Bài toán sẽ trở về 2 chiều. Khi cần dự đoán những chỉ số đánh giá của người dùng nào đó đối với đối tượng nào đó trong điều kiện ngữ cảnh là thời gian 1, ta dùng khối thứ nhất. Tương tự với khối thứ 2, v.v… |
| Slide 15, 16: PP thu giảm số chiều – thuật toán (2 phút) |
| Blah blah blah …. |
| Slide 17: Vào phần hiện thực hóa (2s) |
| Sau đây sẽ là phần hiện thực hóa. |
| Slide 18: mô hình tổng quan (1 phút) |
| Hệ thống gồm các thành phần sau: trình bày sơ mỗi phần có vai trò gì, làm gì …  Hệ khuyến nghị: …  Cơ sở dữ liệu giao tác:  Kho dữ liệu:  OLAP cube:  Dịch vụ web WCF:  Điện thoại Android: … |
| Slide 19: Các điều kiện ngữ cảnh đc chọn (45s) |
| * Nhóm khảo sát 100 người với câu hỏi “khi du lịch, điều kiện ngữ cảnh nào ảnh hưởng đến quyết định chọn lựa điểm du lịch của bạn nhất?” * kết quả như trên hình: với 4 điều kiện được người dùng quan tâm nhất là Bạn đồng hành, kinh phí, thời tiết, thời gian (nhắc đến cấu trúc phân cấp) |
| Slide 20: Hiện thực hóa, demo (5 phút) |
| Sau đây, nhóm xin tiến hành demo ứng dụng Android sử dụng hệ khuyến nghị.  Blah blah blah theo kịch bản cũ, có vần đề gì thì show đoạn clip 5 phút. |
| Slide 21: Vào phần Đánh giá và kết luận (2s) |
| Sau đây là phần đánh giá và kết luận. |
| Slide 22: Movielens (1 phút) |
| Để đánh giá thuật toán trên không gian khuyến nghị 2 chiều, nhóm tác giả dùng bộ dữ liệu Movielens 100.000 dòng, không chứa các điều kiện ngữ cảnh. Bộ này là những chỉ số đánh giá của những người dùng đối với những bộ phim khác nhau được nhóm nghiên cứu GroupLen bên Mỹ thu thập.   * Chỉ số dùng đánh giá là MAE … * Kỹ thuật thực hiện là Cross Validation với 10 lần chạy. Chia dữ liệu 10 phần, 1 phần test, 9 phần train. Lần 1 phần 1 test, còn lại train, lần 2, lần 3 … Sau 10 lần, tính được chỉ số MAE trung bình là 0.7578 sát với các nghiên cứu 2 chiều trước đây. |
| Slide 23: dữ liệu thực tế (45s) |
| Nhóm tác giả xây dựng 1 trang web nhỏ để thu thập dữ liệu, Kết quả được 820 dòng từ 178 người dùng đánh giá trong các điều kiện ngữ cảnh khác nhau.  Sau đây là hình thống kế các chỉ số đánh giá ứng với mỗi điều kiện ngữ cảnh:  Ta thấy cho thấy các chỉ số đánh giá không trải đều ở các điều kiện ngữ cảnh khác nhau. Có những điều kiện ngữ cảnh được nhiều người dùng quan tâm chọn lựa, cũng có những điều kiện ngữ cảnh người dùng ít chọn đến khi muốn đi du lịch. Qua đó cho thấy rõ ràng những điều kiện ngữ cảnh có ảnh hưởng đến việc người dùng chọn lựa địa điểm du lịch, và đương nhiên cũng ảnh hưởng đến mức độ thích thú, chỉ số đánh giá dành cho địa điểm du lịch đó. |
| Slide 24: các thông số đánh giá (45s) |
| Để đánh giá nhận xét kết quả, nhóm dùng:  Chỉ số MAE kèm theo một vài hệ số nữa là:  Hệ số biến thiên, số lượng đánh giá trung bình cho 1 địa điểm và tống số lượng đánh giá (tính theo từng phân khúc dữ liệu) |
| Slide 25: show bảng (đứng nhận xét) (1 phút) |
| Slide 26: đánh giá chung (1 phút) |
| Slide 27: hướng phát triển (1 phút) |