1. Giới thiệu bài toán

* Xây dựng ứng dụng hỗ trợ khách du lịch tại Tp.HCM.
* Hệ thống khuyến nghị các điểm du lịch dựa trên ngữ cảnh(thời gian, thời tiết, bạn đồng hành, tâm trạng, …).

1. Hướng tiếp cận

* Phát triển từ các hệ thống khuyến nghị truyền thống(2 chiều: User x Item), mở rộng lên với nhiều chiều ngữ cảnh khác nhau sử dụng Data warehouse/OLAP với các ưu điểm về:
* Hỗ trợ đa chiều.
* Khả năng kết hợp phân cấp (chẳng hạn một điểm du lịch thích hợp đi vào buổi tối – Thời gian cấp 1, mùa hè – Thời gian cấp 2).
* Khả năng khai thác thông tin của người dùng (ví dụ: sở thích, thói quen), đối tượng khuyến nghị (ví dụ: địa điểm lịch sử, điểm vui chơi giải trí, góc thư giãn, …) và các ngữ cảnh khác.
* Ước tính điểm đánh giá (bài toán nhiều chiều) bằng cách kết hợp phương pháp giảm số chiều và CF truyền thống (2 chiều):

Định nghĩa segment: một segment là một tập con của dữ liệu đánh giá trước đó. Ví dụ: Segment mùa hè, Segment đi cùng bạn bè, Segment trời mưa và buổi tối…

Ý tưởng chính của thuật toán là tìm ra segment phù hợp nhất với điều kiện ngữ cảnh cần xét đồng thời có độ chính xác tốt nhất khi áp dụng thuật toán 2 chiều (user x item) trên các segment đó.

* Phương pháp so sánh thuật toán trên các segment: Chia tập dữ liệu ban đánh giá đang có thành 2 tập: tập dữ liệu huấn luyện (1) và tập dữ liệu đánh giá (2). Thực hiện tính toán đánh giá lại với điều kiện ngữ cảnh của tập (2) trên tập (1) và so sánh độ lệch với thực tế đã có.

Thuật toán gồm 2 pha: pha thứ nhất (chạy offline) tìm ra thứ tự các segment có độ chính xác giảm dần và cao hơn tập dữ liệu gốc (Full segment); pha 2 (chạy online) tìm ra segment trên các segment với độ ưu tiên ở pha 1phù hợp với ngữ cảnh cần xét và thực hiện tính toán với thuật toán CF 2 chiều trên segment đó.

1. Hiện thực bài toán



Gồm có 2 phần chính:

* Phần 1 (Front-end) là ứng dụng trên nền tảng Android hỗ trợ tìm kiếm thông tin các điểm du lịch, tận dụng GPS và Google Maps để truyền tải thông tin tốt nhất.
* Phần 2 (Back-end) là hệ thống khuyến nghị người dùng dựa trên ngữ cảnh với cơ sở dữ liệu, Data warehouse và OLAP Cube.