**CHƯƠNG 4: PHƯƠNG PHÁP KHUYẾN NGHỊ ÁP DỤNG TRONG KHÓA LUẬN**

Sau khi tiến hành khảo sát các phương pháp khuyến nghị ở những phần trên, ở phần này chúng tôi sẽ trình bày chi tiết về phương pháp đã được hiện thực trong khóa luận này. Chương này gồm có 2 phần: phần 1 là hướng tiếp cận cắt giảm số chiều (Reduction-based), phần 2 sẽ trình bày về mô hình hồi qui được sử dụng kết hợp để dự đoán đánh giá người dùng.

**4.1. Hướng tiếp cận cắt giảm số chiều**

Ý tưởng chính của phương pháp được chúng tôi lựa chọn này là đưa bài toán đa chiều về bài toán hai chiều (User × Item) đã có lời giải. Do đó một trong những ưu điểm của phương pháp này là có thể kết hợp với bất kỳ một phương pháp khuyến nghị hai chiều nào đã trình bày ở phần 2.1 sau khi bước giảm chiều hoàn tất. Để hiểu phương pháp này chúng ta cùng bắt đầu quay lại với bài toán khuyến nghị hai chiều với ví dụ sau:

Với hàm dự đoán đánh giá trên, ta có D là tập chứa bộ dữ liệu <user, item, rating> cho mỗi đánh giá của từng người dùng với một đối tượng cụ thể, chúng ta có thể dự đoán bất kỳ một đánh giá nào. Tương tự ta thêm một chiều mới là thời gian thì sẽ có công thức dự đoán sẽ là:

Ở đây D là tập chứa bộ dữ liệu <user, item, time, rating> cho mỗi đánh giá cụ thể. Hàm dự đoán đánh giá trong không gian 3 chiều trên có thể biểu diễn thông qua hàm dự đoán 2 chiều như sau:

Trong đó D[Time = t](User, Item, Rating) là tập dữ liệu đánh giá trích xuất từ tập D bằng cách chọn ra những dữ liệu mà chiều Time có giá trị là t và sau đó chỉ giữ lại trường User, Item, và Rating. Nói cách khác, nếu ta biểu diễn tập dữ liệu D trong cơ sở dữ liệu quan hệ, thì D[Time = t](User, Item, Rating) là tập dữ liệu chọn lọc từ D bằng hai phép toán quan hệ: phép chọn trước và phép chiếu sau.

Cần lưu ý trong nhiều trường hợp thì tập D[Time = t](User, Item, Rating) có thể không có đủ dữ liệu đánh giá để dự đoán bằng thuật toán dự đoán hai chiều. Do đó, phương pháp tổng quát để cắt giảm số trong hệ thống khuyến nghị đa chiều có thể sẽ không sử dụng chính xác ngữ cảnh t để dự đoán đánh giá (u, i, t) mà thay vào đó là tập ngữ cảnh St (contextual segment) chứa các ngữ cảnh cao hơn hoặc bằng với t.