

Bài tập thực hành: Khai thác tập phổ biến & Luật kết hợp

Bộ môn: KHMT

Thời gian làm bài: 1 tuần (Xem deadline trong link nộp bài trên moodle)

Nộp bài:

- Nộp bài lên moodle.
- Đặt tên bài nộp theo định dạng MSSV.rar. Trong đó bao gồm:
 - Tập tin báo cáo.
 - Các tập tin dữ liệu theo yêu cầu của bài tập

Các hành vi sử dụng toàn bộ/một phần bài làm của người khác sẽ bị 0 điểm cho toàn bộ phần thực hành

Bài 1. Apriori

Cho cơ sở dữ liệu giao dịch sau, với minsup = 35% và minconf = 80%

Mã giao dịch	Hạng mục
1	Bread, Peanuts, Milk, Fruit, Jam
2	Bread, Jam, Soda, Chips, Milk, Fruit
3	Steak, Jam, Soda, Chips, Bread
4	Jam, Peanuts, Milk, Fruit
5	Jam, Soda, Chips, Milk, Bread
6	Fruit, Chips, Milk, Yogurt
7	Fruit, Soda, Peanuts, Milk
8	Fruit, Peanuts, Cheese, Yogurt

1. Hãy sử dụng thuật toán Apriori trong công cụ WEKA để tìm tất cả các tập phổ biến thỏa minsup đã cho. Yêu cầu:
 - a. Trình bày các tập phổ biến tìm được theo kích thước tập hạng mục.
 - b. Chụp màn hình WEKA lúc thiết lập tham số cho thuật toán và màn hình kết quả các tập phổ biến.
 - c. Nộp lại tập tin Food.ARFF biểu diễn cơ sở dữ liệu giao dịch nói trên.
2. Hãy sử dụng thuật toán Apriori trong WEKA tìm các luật kết hợp thỏa ngưỡng minsup và minconf đã cho. Yêu cầu:
 - a. Trình bày các luật tìm được theo từng tập hạng mục.
 - b. Chụp màn hình kết quả luật kết hợp.