

Bài tập 1. Bạn đang thực hiện một nhiệm vụ phân loại nhị phân để dự đoán liệu một bệnh nhân có mắc bệnh tim (Có hoặc Không) dựa trên một số đặc điểm. Xác định (các) đặc điểm nào (Tuổi, Huyết áp, Cholesterol, Người hút thuốc) đóng góp nhiều nhất vào việc dự đoán bệnh tim trong trường hợp này tập dữ liệu sử dụng Information Gain.

ID bệnh nhân | Tuổi | Huyết Áp | Cholesterol | Người hút thuốc | Bệnh tim

- 1 Trẻ Cao Bình thường Có Có
- 2 tuổi Cao Cao Không Không
- 3 Trung bình Bình thường Có Có
- 4 Cũ Bình thường Cao Không Không
- 5 Trẻ Cao Cao Có Có
- 6 tuổi Bình thường Cao Không Không
- 7 Trung bình Bình thường Có Có

- 8 Trung Trung Cao Không Không