

BỘ MÔN HỆ THỐNG THÔNG TIN – KHOA CÔNG NGHỆ THÔNG TIN

ĐẠI HỌC KHOA HỌC TỰ NHIÊN THÀNH PHỐ HỒ CHÍ

TRỰC QUAN HÓA DỮ LIỆU ĐỒ ÁN THỰC HÀNH

GIAI ĐOẠN 2: TASK ABSTRACTION



HỌC KỲ II – NĂM HỌC 2022-2023

1. Thông tin nhóm:

STT	Mã sinh viên	Họ và tên
1	19120216	Nguyễn Thụy Ngọc Hân
2	19120453	Nguyễn Dương Gia Bân
3	19120733	Lê Hoàng Thịnh Như Ý
4	19120654	Nguyễn Đức Thắng
5	19120636	Nguyễn Phước Sâm

2. Nội dung báo cáo:

- a. **Task 1:** Thống kê tổng số giao dịch của khách hàng theo loại thẻ mà khách hàng nắm giữ
 - Lấy danh sách khách hàng có loại thẻ là “Blue”
 - ◆ Action: Search > Browse
 - Tính tổng số khách hàng có loại thẻ là “Blue”
 - ◆ Action: Analyze > Produce > Derive
 - Tính tổng số giao dịch của khách hàng theo loại thẻ “Blue”
 - ◆ Action: Analyze > Produce > Derive
 - Lấy danh sách khách hàng có loại thẻ là “Gold”
 - ◆ Action: Search > Browse
 - Tính tổng số khách hàng có loại thẻ là “Gold”
 - ◆ Action: Analyze > Produce > Derive
 - Tính tổng số giao dịch của khách hàng theo loại thẻ “Gold”
 - ◆ Action: Analyze > Produce > Derive

- Lấy danh sách khách hàng có loại thẻ là “Silver”
 - ◆ Action: Search > Browse
- Tính tổng số khách hàng có loại thẻ là “Silver”
 - ◆ Action: Analyze > Produce > Derive
- Tính tổng số giao dịch của khách hàng theo loại thẻ “Silver”
 - ◆ Action: Analyze > Produce > Derive
- Lấy danh sách khách hàng có loại thẻ là “Platinum”
 - ◆ Action: Search > Browse
- Tính tổng số khách hàng có loại thẻ là “Platinum”
 - ◆ Action: Analyze > Produce > Derive
- Tính tổng số giao dịch của khách hàng theo loại thẻ “Platinum”
 - ◆ Action: Analyze > Produce > Derive
- Action: Query > Summarize
- Target: Attribute > Distribution

b. Task 2: Thống kê tỉ lệ khách hàng quay lại của khách hàng dựa trên phân loại Naive Bayes 1, 2 và tổng số tiền giao dịch của khách hàng

- Phân loại khả năng quay lại của khách hàng dựa trên giá trị của Naive Bayes, chia thành 4 nhóm: Rất cao (≥ 0.8), Cao ($0.7 \rightarrow 0.8$), Trung bình ($0.4 \rightarrow 0.7$), Thấp (< 0.4)
 - ◆ Action: Analyze > Produce > Derive
- Lấy danh sách khách hàng có khả năng quay lại “Rất cao”
 - ◆ Action: Search > Browse
- Tính tổng số tiền giao dịch của khách hàng có khả năng quay lại “Rất cao”
 - ◆ Action: Analyze > Produce > Derive
- Lấy danh sách khách hàng có khả năng quay lại “Cao”
 - ◆ Action: Search > Browse
- Tính tổng số tiền giao dịch của khách hàng có khả năng quay lại “Cao”
 - ◆ Action: Analyze > Produce > Derive
- Lấy danh sách khách hàng có khả năng quay lại “Trung bình”
 - ◆ Action: Search > Browse
- Tính tổng số tiền giao dịch của khách hàng có khả năng quay lại “Trung bình”
 - ◆ Action: Analyze > Produce > Derive
- Lấy danh sách khách hàng có khả năng quay lại “Thấp”
 - ◆ Action: Search > Browse
- Tính tổng số tiền giao dịch của khách hàng có khả năng quay lại “Thấp”
 - ◆ Action: Analyze > Produce > Derive

c. Task 3: Thống kê số lượng khách hàng ở các độ tuổi

→ Đếm số khách hàng ở từng độ tuổi

◆ Action: Analyze > Produce > Derive

→ Action: Query > Summarize

→ Target: Attribute > Distribution

d. Task 4: Thống kê số lượng khách hàng nam/nữ của mỗi loại thẻ mà khách hàng nắm giữ

-> Xác định khách hàng nam/ nữ sử dụng các loại thẻ

Action:

- Analyze > product > Present

- Search -> Browser

-> Đếm số lượng khách hàng nam/ nữ sử dụng từng loại thẻ

Action:

- Analyze -> Produce -> Derive

→ Target: Attributes > distributes

e. Task 5: Tỷ lệ khách hàng dùng thẻ 'gold' ở các mức thu nhập

-> Xác định khách hàng sử dụng thẻ gold ở các mức thu nhập khác nhau

Action:

- Analyze > product > Present

- Search > Browser

-> Đếm số lượng khách hàng sử dụng thẻ gold ở các mức thu nhập khác nhau

Action:

- Analyze -> Produce -> Derive

-> Đếm số lượng khách hàng sử dụng thẻ gold

Action:

- Analyze -> Produce -> Derive

-> Tính tỷ lệ khách hàng sử dụng thẻ "gold" ở các mức thu nhập khác nhau

Action:

- Analyze -> Produce -> Derive

→ Target: Attributes > distributes

f. Task 6: Thống kê số lượng khách hàng có khả năng rời đi / không rời đi khi số tháng không hoạt động trên 6 tháng

-> Xác định danh sách khách hàng có khả năng rời đi/ không rời đi khi số tháng không hoạt động trên 6 tháng

Action:

- Analyze > product > Present

- Search > Browser

-> Đếm số lượng khách hàng có khả năng rời đi/ không rời đi khi số tháng không hoạt động trên 6 tháng

- Analyze -> Produce -> Derive

→ Target: Attributes > distributes

g.Task 7: Tìm loại thẻ khách hàng mà khách hàng nắm giữ nhiều nhất

-> Xác định khách hàng nắm giữ từng loại thẻ

Action:

- Analyze > product > Present

- Search > Browser

- Query > compare

-> Đếm số khách hàng nắm giữ từng loại thẻ

Action:

- Analyze -> Produce -> Derive

→ Target: Attributes > extremes

