

# THÀNH VIÊN NHÓM

Tên môn học: Nhập Môn Công Nghệ Tri Thức Và Máy học

Mã Môn : CS110.I11

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| MSSV | HỌ TÊN | EMAIL | ĐIỆN THOẠI | CHỮ KÍ |
| 14520401 | Trương Ngọc Kha | [14520401@gm.uit.edu.vn](mailto:14520401@gm.uit.edu.vn) |  |  |
| 14520422 | Huỳnh Đăng Khoa | [14520422@gm.uit.edu.vn](mailto:1312447@student.fit.hcmus.edu.vn) |  |  |
| 14520431 | Trương Đăng Khoa | [14520431@gm.uit.edu.vn](mailto:14520431@gm.uit.edu.vn) |  |  |

1. **Source code và video demo :**

Link download : <https://drive.google.com/file/d/1nv0iwW0xg-axho043arbatcqKHrKWCmH/view?usp=sharing>

1. **Nội dung:**
2. **Giới thiệu :**

* Cây quyết định là một kiểu mô hình dự báo
* Kỹ thuật học máy dùng trong cây quyết định được gọi là học bằng cây quyết định, hay chỉ gọi với cái tên ngắn gọn là cây quyết định
* Phương tiện có tính mô tả dành cho việc tính toán các xác suất có điều kiện
* Sự kết hợp của các kỹ thuật toán học và tính toán nhằm hỗ trợ việc mô tả, phân loại và tổng quát hóa một tập dữ liệu cho trước
* Cây quyết định là một cấu trúc phân cấp của các nút và các nhánh
* 3 loại nút trên cây:
* Nút gốc
* Nút nội bộ: mang tên thuộc tính của CSDL
* Nút lá: mang tên lớp C*i*
* Nhánh: mang giá trị có thể của thuộc tính
* Cây quyết định được sử dụng trong phân lớp bằng cách duyệt từ nút gốc của cây cho đến khi đụng đến nút lá, từ đó rút ra lớp của đối tượng cần xét

1. **Ví dụ :**

* David là quản lý của một câu lạc bộ đánh golf nổi tiếng. Anh ta đang có rắc rối chuyện các thành viên đến hay không đến. Có ngày ai cũng muốn chơi golf nhưng số nhân viên câu lạc bộ lại không đủ phục vụ. Có hôm, không hiểu vì lý do gì mà chẳng ai đến chơi, và câu lạc bộ lại thừa nhân viên.
* Mục tiêu của David là tối ưu hóa số nhân viên phục vụ mỗi ngày bằng cách dựa theo thông tin dự báo thời tiết để đoán xem khi nào người ta sẽ đến chơi golf. Để thực hiện điều đó, anh cần hiểu được tại sao khách hàng quyết định chơi và tìm hiểu xem có cách giải thích nào cho việc đó hay không.
* Vậy là trong hai tuần, anh ta thu thập thông tin về: Trời (outlook) (nắng (sunny), nhiều mây (overcast) hoặc mưa (raining)). Nhiệt độ (temperature) bằng độ F. Độ ẩm (humidity). Có gió mạnh (wind) hay không.
* Và tất nhiên là số người đến chơi golf vào hôm đó. David thu được một bộ dữ liệu gồm 14 dòng và 5 cột.

































