



fit@hcmus

TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN, ĐHQG - HCM
KHOA CÔNG NGHỆ THÔNG TIN

MÔN HỌC: MÁY HỌC

**BÀI TẬP THỰC HÀNH II
PHÂN LỚP
BẰNG MẠNG NƠ RON**

1 Nội dung bài tập

1.1 Giới thiệu bài toán

Trong nông nghiệp, năng suất của cây trồng phụ thuộc rất nhiều vào thời tiết, chế độ chăm sóc, sâu bệnh,... để tăng năng suất người trồng cần có những biện pháp hạn chế những tác động của các yếu tố trên.

Tuy nhiên, bản thân cây trồng cũng có những loại bệnh nhất định và thường biểu hiện ra bên ngoài trên thân, lá... những biểu hiện đó bất thường trên cây trồng thường liên quan tới một loại bệnh nào đó.

Nhằm giúp người trồng xác định được loại bệnh đang có trên cây trồng, sinh viên được cung cấp tập dữ liệu hình ảnh về các loại bệnh trên cây và tập dữ liệu đã được xác định được loại bệnh (train.csv và test.csv) bao gồm 4 loại:

- Combinations
- Healthy
- Rust
- Scab

Datasets [plant-pathology-2020-fgvc7.zip - Google Drive](#)

Xây dựng mô hình mạng nơ-ron để xác định loại bệnh cho một ảnh đầu vào bất kỳ.

1.2 Mục tiêu

Tìm hiểu và vận dụng các kỹ thuật:

- **Nhận diện bài toán:** Sinh viên cần tìm hiểu bài toán và dữ liệu được giao nhằm xác định nội dung và ý nghĩa bài toán thực tế cần giải quyết. Thông qua đó, sinh viên có khả năng ánh xạ vấn đề thực tế sang bài toán lập trình.
- **Giải quyết vấn đề:** sinh viên được yêu cầu đưa ra các giải pháp và hướng tiếp cận nhằm giải quyết được yêu cầu bài toán thực tế.
- **Xử lý và phân tích dữ liệu:** sinh viên có khả năng xử lý các công cụ phân tích dữ liệu tự động nhằm tìm ra các thông tin hữu ích, các đặc trưng tiềm ẩn ảnh hưởng đến mục tiêu bài toán.
- **Thiết kế và cài đặt các thuật toán đã học:** sinh viên được yêu cầu có khả năng đề xuất, triển khai và giải thích các thuật toán mạng Nơron đã học nhằm giải quyết bài toán được giao.

1.3 Yêu cầu

Nhóm theo danh sách nhóm đã đăng ký và phải thực hiện các yêu cầu tối thiểu sau:

1. Phân tích kỹ bài toán và tập dữ liệu hình ảnh được cung cấp. Chọn lựa và trình bày kiểu mạng nơron để giải quyết bài toán.
2. Cài đặt mạng Nơron.
3. Báo cáo kết quả đạt được sau quá trình phân tích và cài đặt.

Lưu ý:

- Nhóm có thể thực hiện thêm nhiều phân tích và các mạng Nơron khác. Dựa trên kết quả báo cáo và phân tích, nhóm sẽ được tính điểm cộng tối đa 20% điểm của bài tập này. Điểm tối đa của bài tập là min(điểm bài tập + điểm cộng, 10).
- Báo cáo cần có bảng đánh giá đóng góp và tự đánh giá của các thành viên trong nhóm.
- Các thuật toán phải được cài đặt bằng ngôn ngữ **Python** trên môi trường **Google Colab**.

2 Quy định nộp bài

Các nhóm cần nộp 1 file nén duy nhất có tên

<MSSV1_MSSV2_MSSV3>.zip/rar (ví dụ: 1912000_1912001_1912003.zip) trong đó chứa:

- Báo cáo: dưới dạng **pdf** chứa nội dung phân tích, các số liệu, dẫn chứng, và thuật toán nhóm đã cài đặt.
- Mã nguồn: tập tin **.py** hoặc **.ipynb**. Lưu ý mã nguồn phải có khả năng chạy được (runable) trên môi trường colab đã được quy định ở mục 1.3.