

HW01: Pattern Recognition

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

Yêu cầu

Sinh viên tự thực hiện các yêu cầu sau:

1. Phân tích kỹ bài toán và tập dữ liệu hình ảnh được cung cấp. Lựa chọn một thuật toán phân lớp đã học LDA, SVM để giải quyết bài toán.
2. Báo cáo kết quả đạt được sau quá trình phân tích và cài đặt.

Lưu ý về việc trình bày

- Sinh viên có thể thực hiện thêm nhiều phân tích và các thuật toán khác. Dựa trên kết quả báo cáo và phân tích, sinh viên 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).
- Các thuật toán phải được cài đặt bằng ngôn ngữ **Python** trên môi trường [Google Colab](#).