

YÊU CẦU NỘP PROJECT MÔN HỌC MÁY

Hạn nộp

Hạn chót: 24h00 ngày 12 tháng 01 năm 2026

1. Hình thức nộp

- Sinh viên nộp **link GitHub** chứa toàn bộ project.
- Thư mục GitHub cần được **tổ chức khoa học**, dễ đọc, dễ chạy.
- Các thành phần bắt buộc trong repo:

```
📦 <project_name>/  
    └── README.md           # Hướng dẫn cài đặt và chạy project  
    └── Nx_report.pdf       # Báo cáo project (file PDF), đặt tên dưới dạng Nx_report.pdf  
      (trong đó x là số nhóm)  
    └── data/                # Thư mục chứa dữ liệu (hoặc link tải data nếu file lớn)  
    └── src/                 # Mã nguồn Python, tách thành các file/hàm rõ ràng  
      └── preprocessing.py   # Xử lý dữ liệu  
      └── eda.py            # Phân tích dữ liệu khám phá (EDA)  
      └── feature_engineering.py # Tạo và chọn đặc trưng (Nếu có)  
      └── model_hovaten.py    # Huấn luyện mô hình, đặt tên dưới dạng model_hovaten.py,  
      (trong đó model là tên mô hình, hovaten thay bằng họ và tên người thực hiện)  
      └── evaluation.py      # Đánh giá mô hình  
      └── app.py / main.py    # Web app hoặc script chính để chạy demo (Tùy chọn)  
    └── requirements.txt     # Danh sách thư viện cần cài  
    └── web/                 # Code giao diện web hoặc dashboard (Tùy chọn)
```

Có thể thêm file nếu cần

2. Yêu cầu về mã nguồn

- Code chia theo pipeline chuẩn (preprocessing → EDA → model → evaluation)
- Không gộp tất cả vào 1 file
- Mỗi hàm có comment mô tả rõ ràng

3. Yêu cầu về phần Web/Dashboard (Nếu có)

- Có giao diện web hoặc dashboard hiển thị kết quả mô hình
- Web phải cho phép:
 - Tải dữ liệu hoặc chọn file mẫu
 - Hiển thị biểu đồ phân tích
 - Cho phép nhập input và hiển thị dự đoán

4. Báo cáo (Report PDF)

Báo cáo phải trình bày theo cấu trúc khoa học, gồm:

1. Giới thiệu đề tài
2. Mục tiêu và bài toán đặt ra
3. Mô tả dữ liệu và các bước tiền xử lý
4. Mô hình học máy sử dụng (Mô tả ngắn gọn nguyên lý hoạt động của mô hình đó + Lý do lựa chọn mô hình + Train model + Hyperparameter tuning). Nếu chỉ fit đơn giản điểm phần mô hình sẽ chỉ đạt mức tối thiểu.
5. Kết quả và đánh giá mô hình (độ chính xác, biểu đồ, confusion matrix, precision/recall/f1-score, roc-auc, cross-validation – Đánh giá chi tiết)
6. Kết luận và hướng phát triển
7. Tài liệu tham khảo

5. Điểm cộng

- Phát triển ứng dụng web có giao diện thân thiện, dễ sử dụng
- Có biểu đồ trực quan hóa sinh động
- So sánh nhiều mô hình khác nhau
- Áp dụng kỹ thuật nâng cao