

Phân công công việc nhóm - Dự án phân loại thư rác bằng Naive Bayes

1. Thành viên

- Hồ Công Chí (Lead)
- Nguyễn Đoàn Thanh Thảo
- Vũ Huỳnh Phát

2. Mục tiêu dự án

Phát triển một hệ thống phát hiện và phân loại thư rác (spam email classification) hiệu quả, sử dụng thuật toán Naive Bayes – từ xử lý dữ liệu đến huấn luyện, triển khai, và đánh giá mô hình.

3. Phân công công việc chi tiết

Hồ Công Chí (Leader)

- Thiết kế kiến trúc giải pháp tổng thể (end-to-end workflow).
- Xây dựng pipeline huấn luyện mô hình Naive Bayes.
- Phân tích dữ liệu đầu ra, đưa ra giải pháp, đề xuất cải tiến mô hình.
- Viết báo cáo mô hình và kiểm duyệt tiến độ nhóm.

Nguyễn Đoàn Thanh Thảo

- Thu thập, làm sạch và chuẩn hóa dữ liệu
 - Trực quan hóa dữ liệu
 - Thực hiện tiền xử lý: Làm sạch, tách từ, vector hóa văn bản...
- Result: Dataset sạch, báo cáo thống kê dữ liệu, pipeline xử lý.

Vũ Huỳnh Phát

- Đánh giá hiệu suất mô hình
 - Tối ưu hóa pipeline và hiệu năng mô hình.
 - Demo
- Result: Notebook đánh giá mô hình, báo cáo performance kỹ thuật.

4. Kế hoạch thời gian

- Tuần 1: Thu thập, chuẩn hóa và phân tích dữ liệu – Ken
- Tuần 2: Xây dựng mô hình Naive Bayes và huấn luyện – Chí
- Tuần 3: Đánh giá mô hình, điều chỉnh tham số – Phát

- Tuần 4: Cải tiến mô hình, cập nhật dữ liệu, hoàn thành tiểu luận – Cả 3

5. Ghi chú

- Mỗi người viết báo cáo phần mình phụ trách.
- Họp cuối mỗi tuần trong zalo để cập nhật tiến độ và điều chỉnh nhé.