

KẾ HOẠCH LÀM VIỆC VÀ PHÂN CÔNG NHIỆM VỤ

Dự án: Phân tích Xu hướng Âm nhạc Toàn cầu trên Spotify (2009-2025)

1 Thông tin thành viên (Team Member Information)

Nhóm bao gồm 02 thành viên với vai trò và thông tin như sau:

STT	Họ và Tên	MSSV	Vai trò chính
1	Lê Hoàng Mỹ Hạ	23120038	Data Analyst, Modeler
2	Nguyễn Mạnh Thắng	23120084	Data Engineer

Bảng 1: Danh sách thành viên nhóm

2 Phân công công việc (Work Distribution)

Dưới đây là bảng chi tiết các hạng mục công việc, người phụ trách và tỷ lệ đóng góp:

Giai đoạn	Nhiệm vụ cụ thể	Người làm	Đóng góp
1. Data Exploration	Thu thập dữ liệu, phân tích cấu trúc	Mỹ Hạ	100%
	Vẽ biểu đồ EDA (Hist, Heatmap)	Mạnh Thắng	100%
2. Preprocessing	Xử lý Missing Value, Duplicates	Mỹ Hạ	50% / 50%
	Feature Engineering (Genre, Log Transform)	Mỹ Hạ	100%
	Encoding (One-hot) & Scaling	Mạnh Thắng	100%
3. Modeling	Chia tập Train/Test, Xây dựng Model	Mạnh Thắng	100%
	Đánh giá Model (Metrics), Tinh chỉnh	Cả hai	50% / 50%
4. Reporting	Viết báo cáo, Slides, Readme	Cả hai	50% / 50%
TỔNG KẾT ĐÓNG GÓP TOÀN DỰ ÁN			50% - 50%

3 Quy trình hợp tác (Collaboration Process)

Để đảm bảo dự án vận hành trơn tru, nhóm đã thống nhất quy trình làm việc như sau:

3.1 Công cụ hỗ trợ

- **Quản lý mã nguồn (Source Control):** Sử dụng GitHub để lưu trữ code, quản lý phiên bản và merge các phần làm việc của thành viên.
- **Môi trường lập trình:** Jupyter Notebook (Local) và VS Code.
- **Giao tiếp:** Messenger để trao đổi nhanh hàng ngày. Google Meet để họp tiến độ (1 lần/tuần).
- **Quản lý tác vụ:** Sử dụng Notion/Trello để theo dõi Deadline.

3.2 Phương pháp làm việc nhóm

1. **Chia để trị (Divide and Conquer):** Dự án được chia nhỏ thành các module độc lập (EDA, Preprocessing, Modeling) để hai thành viên có thể làm việc song song mà không bị xung đột code.
2. **Code Review:** Trước khi merge vào nhánh chính ('main'), code của thành viên này phải được thành viên kia xem qua (pull request) để phát hiện lỗi.
3. **Giải quyết xung đột:** Nếu có bất đồng về hướng giải quyết (ví dụ: cách xử lý Outliers), nhóm sẽ cùng chạy thử nghiệm cả hai phương án và chọn phương án có kết quả tốt hơn trên tập Validation.

4 Kế hoạch và Tiến độ (Project Planning & Timeline)

Dự án được thực hiện trong vòng 4 tuần với các mốc thời gian cụ thể:

Tuần	Giai đoạn	Chi tiết hoạt động	Trạng thái
Tuần 1 (22/11-29/11)	Khởi động	- Tìm kiếm dataset Spotify - Thiết lập Github Repo - Phân tích sơ bộ (EDA)	Hoàn thành
Tuần 2 (30/11-6/11)	Xử lý dữ liệu	- Xử lý Missing Values (Genre) - Log Transform, One-hot Encoding - Chuẩn hóa dữ liệu	Hoàn thành
Tuần 3 (7/12-14/12)	Mô hình hóa	- Train/Test Split - Chạy các mô hình hồi quy - So sánh kết quả MAE/RMSE/R ²	Hoàn thành
Tuần 4 (15/12-22/12)	Tổng kết	- Viết báo cáo cuối kỳ - Hoàn thiện Readme.md	Hoàn thành