**ĐỒ ÁN MÔN HỌC KHAI THÁC DỮ LIỆU**

**LỚP IS252.N21.HTCL**

1. **Tên nhóm:** TEAM 11
2. **Tên đề tài:** PREDICT THE MUSIC GENRE OF A SONG
3. **Tóm tắt nội dung đề tài (không quá 10 dòng):**

Âm nhạc tồn tại khắp mọi nơi và là một yếu tố không thể thiếu trong cuộc sống, nó đem lại ảnh hưởng tích cực đến đời sống tinh thần của con người. Âm nhạc đã phát triển đa dạng hình thức và thể loại theo từng thời kỳ, nhằm phục vụ nhu cầu cho các đối tượng khác nhau.

Trong lĩnh vực truyền thông và truyền phát âm thanh, công ty cần có khả năng tự động phân loại các thể loại âm nhạc để quản lý, sắp xếp dữ liệu và tăng trải nghiệm người dùng. Với sự phát triển của khoa học dữ liệu và các mô hình dự đoán, việc phân loại bài hát đã trở nên quan trọng và cấp thiết với các công ty lớn.

Trong bài nghiên cứu này nhóm đã thực hiện việc phân loại các bài hát thành các thể loại như Jazz, Blues, Electronic, Rap, Hip-Hop,... dựa trên dữ liệu được cung cấp.

1. **Thông tin nhóm:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| STT | MSSV | Họ tên | Mô tả công việc được giao | Tổng công việc được phân (% công việc được phân công) | Khả năng hoàn thanh công việc được giao (% hoàn thanh) | Vai trò | Ký tên |
| 1 | 20521228 | Bùi Đức Duy | - Lựa chọn và đánh giá dữ liệu  - Mô tả dữ liệu, phát biểu bài toán  - Tiền xử lý dữ liệu và trực quan hóa từng thuộc tính  - Random Forest  - Đánh giá thời gian, độ chính xác các thuật toán  - Dự đoán mối quan hệ giữa các thuộc tính đã phân lớp với thuộc tính dự đoán  - Viết báo cáo | 60% | 100% | Nhóm trưởng |  |
| 2 | 20522145 | Nguyễn Thị Cẩm Vân | - Lựa chọn và đánh giá dữ liệu  - Mô tả dữ liệu, phát biểu bài toán  - Naïve Bayes  - Support Vector Machine  - So sánh các thuật toán  -Viết báo cáo | 50% | 100% | Thành viên |  |
| 3 | 20520876 | Hồ Bảo An | - Lựa chọn và đánh giá dữ liệu  - Mô tả dữ liệu, phát biểu bài toán  -K-Nearest Neighbors  - Đánh giá thời gian, độ chính xác các thuật toán | 50% | 100% | Thành viên |  |
| 4 | 20521532 | Lưu Thảo Linh | - Lựa chọn và đánh giá dữ liệu  - Mô tả dữ liệu, phát biểu bài toán  - Decision Tree (CART)  - Support Vector Machine  - So sánh các thuật toán  -Viết báo cáo | 50% | 100% | Thành viên |  |