

# Phân loại thể loại âm nhạc

Nguyễn Đăng Khoa, Lê Thị Trúc Ly, Lê Đoàn Kim Ngân, Lâm Tú Nhi

## Giới thiệu

### Định nghĩa vấn đề

- Input:** Đặc trưng âm thanh, đặc trưng của từng thể loại âm nhạc
- Output:** Thể loại của bản nhạc

### Thách thức

- Dữ liệu chứa giá trị thiếu (NaN), độ dao động cao giữa nghệ sĩ và phong cách, chất lượng âm thanh khác nhau.
- Dataset có độ mất cân bằng: Một số thể loại xuất hiện nhiều hơn hẳn, làm mô hình thiên lệch.

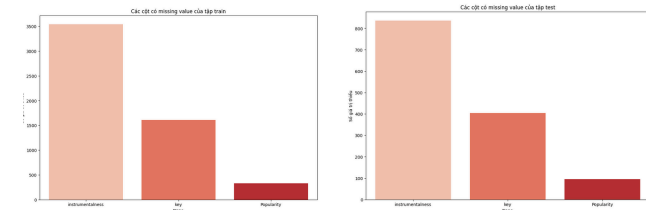
**Mục tiêu :** Xây dựng hệ thống dự đoán giá nhà chính xác, ổn định và có khả năng tổng quát tốt.

## Tập dữ liệu

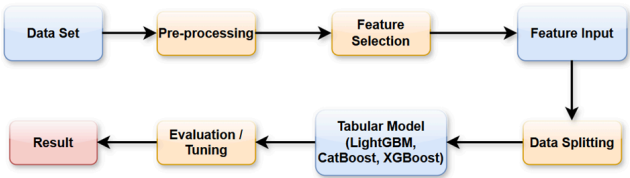
- Nguồn dữ liệu: Music Genre Classification
- Số lượng: 14396 mẫu (training set), 3600 mẫu (test set) và 17 đặc trưng

<https://www.kaggle.com/competitions/shai-music-genre-classification>

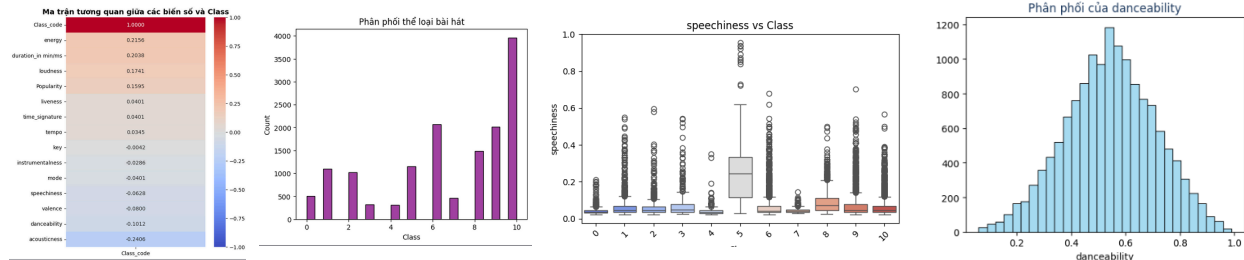
ID	Artist Name	Track Name	Popularity	danceability	energy	key	loudness	mode	speechiness	acousticness	instrumentalness	liveness	valence	tempo	duration_in_ms	time_signature	id3
0	1	Marina Marston	Not Ahead	37.0	0.334	0.536	9.0	-6.649	0	0.0281	0.370000	NaN	0.106	0.235	152.429	204947.0	4
1	2	The Black Keys	Heaven for You	87.0	0.725	0.747	11.0	-5.545	1	0.0576	0.022000	0.0460	0.104	0.380	133.921	191954.0	4
2	3	Royce & The Legend	ghost in	NaN	0.044	0.064	7.0	-6.094	1	0.0919	0.000400	0.0200	0.084	0.401	109.953	140077.0	4
3	4	Detroit Blues Band	Missing You	11.0	0.515	0.306	NaN	-14.711	1	0.0112	0.907000	0.0210	0.300	0.501	170.472	200903.0	3
4	5	Coast Contra	My Lady	48.0	0.565	0.777	8.0	-5.096	0	0.2490	0.180000	NaN	0.211	0.619	88.311	234145.0	4



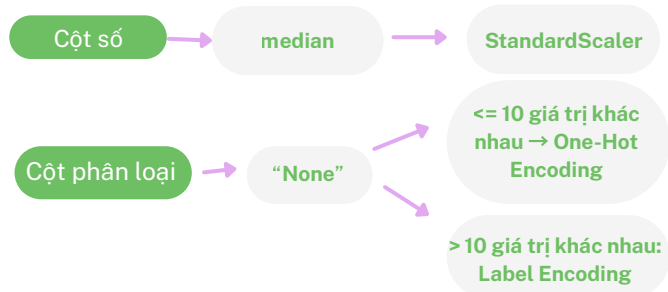
## Phương pháp đề xuất



## 1.Phân tích, khám phá dữ liệu (EDA)



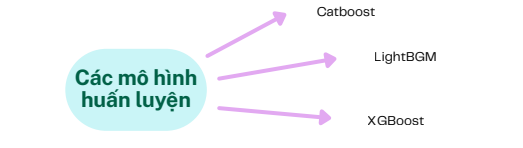
## 2.Tiền xử lí dữ liệu



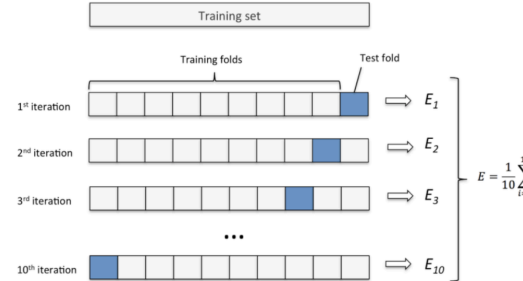
## 3. Kỹ thuật tạo đặc trưng

energy_dance	Nhận biết các bài nhạc có độ "sôi động" cao. (Năng lượng * danceability)	is_loud	Phân biệt nhạc hiện đại (âm lượng cao) với classical hoặc acoustic
energy_valence	Phân biệt nhạc vui và nhạc trầm (năng lượng + cảm xúc tích cực)	duration_minutes	Chuyển thời lượng sang phút
acous_tic_instr	Nhận diện các thể loại như classical, ambient, instrumental (acoustic * instrumental)	energy_dance_ratio	Phân biệt rock/metal (năng lượng cao, dance thấp) với funk/disco (Tỉ lệ năng lượng so với danceability)
tempo_bin	Phân loại nhạc chậm/nhanh rõ ràng hơn.	is_fast	Nhịp nhanh (>120 BPM), phù hợp phân biệt EDM, pop dance với các thể loại chậm.

## 4.Huấn luyện mô hình



## 5. Đánh giá mô hình



KFold chia dữ liệu thành 5 phần bằng nhau , trong đó 4 tập train và 1 tập test

$$Accuracy = \frac{TN + TP}{TN + FP + TP + FN} \quad F1-score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

$$Precision = \frac{TP}{TP + FP}, \quad Recall = \frac{TP}{TP + FN}$$

$$Balanced Accuracy = \frac{1}{K} \sum_{i=1}^K Recall_i = \frac{1}{K} \sum_{i=1}^K \frac{TP_i}{TP_i + FN_i}$$

## Thực nghiệm

- Lần 1:** Baseline: xử lí giá trị thiếu, mã hóa và chuẩn hóa các đặc trưng.
- Lần 2:** Tạo các đặc trưng mới về thời lượng, độ to, nhịp độ,..
- Lần 3:** Kết hợp các mô hình

**Lần 1: Baseline: xử lí giá trị thiếu, mã hóa và chuẩn hóa các đặc trưng.**

Model	accuracy	f1 score	accuracy private	accuracy public
XGBoost	0,559322	0,600757	0,56111	0,56825
CatBoost	0,531259	0,596084	0,52222	0,5369
LightGBM	0,560503	0,58612	0,55648	0,55634

**Lần 2: Tạo các đặc trưng mới về thời lượng, độ to, nhịp độ,..**

Model	accuracy	f1 score	accuracy private	accuracy public
XGBoost	0,557585	0,601	0,56388	0,57301
CatBoost	0,531606	0,58762	0,56111	0,55992
LightGBM	0,55821	0,5907	0,51666	0,54126

**Lần 3: Kết hợp các mô hình**

Model	accuracy	f1 score	accuracy private	accuracy public
Weighted	0,916782	0,936782	0,5324	0,54761

## Kết luận

**Mô hình đạt kết quả trên Kaggle tốt  
→ XGBoost (0,57301)**

	submission_xgb.csv	0.56388	0.57301
Complete (after deadline) ...			