Project Assignment: Deep Learning for Music Genre Classification

Hedi Hadiji

October 2025

1 Overview

The goal of this project is to build and evaluate a deep learning model that classifies music tracks by genre using spectrogram-based representations of audio signals. Students will reproduce a baseline model and then extend it with one meaningful modification of their choice. The emphasis is on understanding data preprocessing, model design, training dynamics, and experimental analysis.

2 Objectives

- Understand how to preprocess audio data for deep learning.
- Implement and train a convolutional neural network (CNN) for classification.
- Investigate the impact of one model or training modification.
- Report and interpret experimental results.

3 Dataset

You will use the **FMA-small** dataset Subset of the Free Music Archive dataset with 8 balanced genres and ∼8,000 tracks.

Link here

Use the **librosa** library to preprocess the data.

4 Baseline Implementation

Each student must first reproduce the following baseline model (using pytorch built-in layers)

- **Input:** Mel-spectrograms with 128 mel bins, normalized per track.
- **Model:** Small (2D) CNN with 3–4 convolutional layers (Conv → ReLU → MaxPool → Dropout), followed by a fully connected layer and softmax output.
- **Loss:** Cross-entropy.
- **Optimizer:** Adam.
- **Metric:** Test accuracy.

5 Extension

Each student must propose and implement one substantial modification. Examples include:

- **Architectural:** Replace or extend the CNN with an RNN, CRNN, or attention layer.
- **Representational:** Compare mel-spectrograms, with other transforms.

- **Training:** Introduce data augmentation (pitch shift, time stretch, noise), label smoothing, or learning-rate scheduling.
- **Regularization:** Experiment with dropout, etc.
- **Transfer learning:** Use embeddings from a different model

State your hypothesis: what changes you are making, why it might help, and how they will measure improvement.

6 Submit

1. **Code:**
   - Baseline reproduction and modified model implementation.
   - Training logs and plots (loss and accuracy curves).
   - Reproducible configuration (e.g., random seeds, environment setup).
2. **Report (max 6 pages):**
   - Dataset description and preprocessing details.
   - Model architecture and modification explanation.
   - Experimental setup, metrics, and results.
   - Quantitative comparison of baseline and modified model.
   - Discussion and interpretation of results.

Deadline: Must be submitted on Edunao by **November 30th**.

7 Grading Rubric

Category	Weight	Description
Technical correctness	25	Baseline reproduced, code runs, and model trains properly.
Experimental rigor	25	Clear, controlled experiments with reproducible results.
Analysis and insight	25	Meaningful interpretation and error analysis.
Report quality	15	Clear writing, organization, and figures.
Creativity / novelty	10	Originality or thoughtful extension.

8 Submission

Submit a single compressed file containing:

- Source code and instructions to run, valid link to an accessible github is possible/good.
- The project report (PDF).