

Temat: Analizator emocji w muzyce

Opis:

Celem projektu jest stworzenie systemu rozpoznającego emocje przekazywane przez utwory muzyczne. Program będzie analizował pliki audio i klasyfikował pod kątem emocjonalnym (zarówno całe utwory, jak i wskazane fragmenty). Plan obejmuje stworzenie modelu opartego o głębokie uczenie, trenowanego na oznaczonych emocjonalnie utworach, przygotowanie danych, eksperymenty w celu poprawy jego skuteczności oraz aplikacji webowej.

Harmonogram:

16 – 22 października 2025

Analiza wymagań i literatury z zakresu MER. Wybór datasetu i przygotowanie środowiska.

23 – 29 października 2025

Implementacja prototypu bazowego ze wstępnie wytrenowanym modelem na danych pozbawionych dodatkowego przetworzenia (czyszczenia i augmentacji). Prototyp umożliwiać będzie wczytanie pliku audio, jego analizę i zwrócenie predykcji.

30 października – 9 listopada 2025

Ekspertyzacje z różnymi architekturami oraz analiza wpływu augmentacji danych. Ocena jakości danych i potencjalne ich wzbogacenie poprzez dobór dodatkowego zbioru danych.

10 – 16 listopada 2025

Kontynuacja treningu modelu i tuning hiperparametrów.

17 – 23 listopada 2025

Ewaluacja na zbiorze testowym i analiza błędów oraz próby ulepszeń.

24 listopada 2025 – 4 grudnia 2025

Stworzenie aplikacji webowej: wczytanie pliku audio, predykcja emocji, wizualizacja wyników. Optymalizacja i przygotowywanie finalnego modelu.

5 listopada – 14 grudnia 2025

Ekspertyzacje uwzględniające analizę wpływu długości utworu, test generalizacji oraz porównanie z modelami klasycznymi. Optymalizacja wydajności systemu.

15 – 21 grudnia 2025

Testy użytkowe, poprawki UX i dopracowanie interfejsu. Przygotowywanie finalnej wersji kodu.

22 grudnia 2025 – 4 stycznia 2026

Finalizacja projektu – przygotowanie dokumentacji, prezentacji, nagrania demo.

Plan eksperymentów:

- Badanie wpływu cech i architektury modelu na skuteczność oraz analiza różnych sposobów reprezentacji utworu. Wyniki przedstawione zostaną w formie tabeli z metrykami (w tym: dokładność, precyzja, czułość, wynik F1) oraz wykresów.
- Badania wpływu zastosowania uczenia transferowego.
- Badania wpływu augmentacji danych oraz wrażliwości na stopień zbalansowania próbek w klasach i technik wyrównywania niezbalansowania licznosci.
- Ocena modelu na różnych długościach utworu w celu poznania jak długość wejściowego fragmentu wpływa na stabilność predykcji. Porównanie, czy emocja jest łatwo rozpoznawana niezależnie od fragmentu oraz spójna z emocją przewidzianą w kontekście całego utworu.
- Porównanie oceny emocji użytkowników dla utworów z etykietami zbioru danych i predykcjami modelu. To pomoże ocenić, na ile rozbieżności modelu wynikają z trudności samego zadania (np. ludzie też się nie zgadzają co do emocji utworu), a na ile z błędu modelu.

Planowana funkcjonalność programu:

- Interfejs użytkownika realizowany poprzez aplikację webową
- Analiza emocji z pliku audio:
 - Użytkownik może wczytać plik muzyczny poprzez interfejs.
 - Program dokonuje analizy i wyświetla wynik w postaci rozpoznanej emocji włącznie ze stopniem pewności.
 - Możliwość wyboru analizowanego fragmentu.
 - Analiza zmian emocji w czasie przy regulowanym oknie czasowym.
- Możliwość porównania skuteczności różnych modeli w zadaniu klasyfikacji emocji w muzyce
- Wizualizacja otrzymywanych predykcji
 - Predykcja uwzględniająca cały kontekst utworu.
 - Predykcja uwzględniająca kontekst zaznaczonego fragmentu utworu.
 - Prezentacja zmiany emocji w czasie.

Stack technologiczny:

- kontrola wersji: git
- język: Python
- biblioteki: pytorch (ML), streamlit (UI), numpy, torchaudio (przetwarzanie audio; przygotowanie danych treningowych), librosa (przetwarzanie audio)
- linter: flake8
- autoformatter: black
- testy automatyczne: pytest, tox (testowanie na różnych wersjach interpretera)
- środowisko wirtualne: venv

Bibliografia

- Liyanarachchi, R., Joshi, A., & Meijering. "A Survey on Multimodal Music Emotion Recognition." *arXiv:2504.18799*, 2025
- Louro, P.L. i in. "A Comparison Study of Deep Learning Methodologies for Music Emotion Recognition." *Sensors*, 24(7):2201, 2024.
- Kang, J., Herremans, D. "Are we there yet? A brief survey of Music Emotion Prediction Datasets, Models and Outstanding Challenges." *arXiv:2406.08809*, 2024