

ДОДАТОК А

Графічний матеріал кваліфікаційної роботи

Харківський національний університет радіоелектроніки

МОДЕЛІ ТА АЛГОРИТМИ КЛАСИФІКАЦІЇ ОБ'ЄКТІВ В МУЗИЧНИХ СТІМІНГОВИХ СИСТЕМАХ ІЗ ВИКОРИСТАННЯМ ГЛИБОКОГО НАВЧАННЯ

Кваліфікаційна робота
другий (магістерський) рівень

Здобувач:
ст.гр.СПм-23-3
Сергородцев І.Д.

Керівник:
проф. каф. ЕОМ
Фесенко Т.Г.

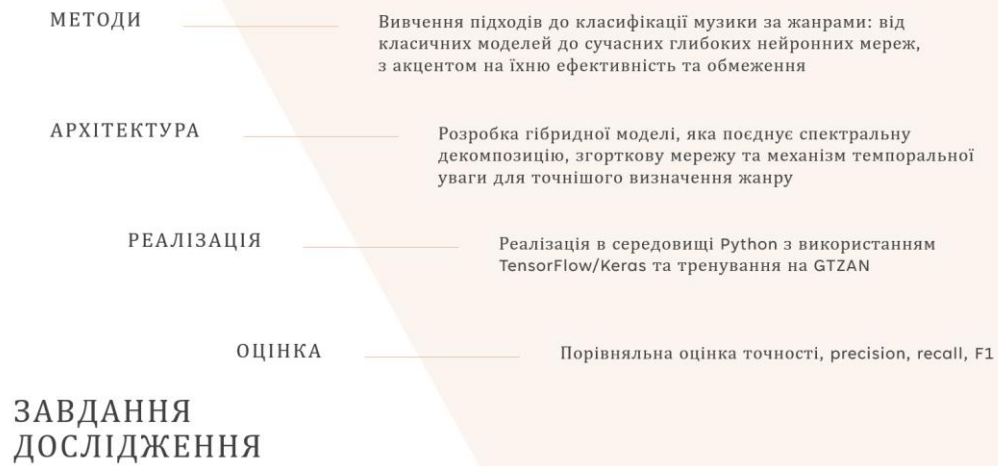
МЕТА ТА АКТУАЛЬНІСТЬ РОБОТИ

Актуальність

- Розвиток потокових сервісів вимагає точного автоматичного аналізу музики.
- Існуючі системи класифікації мають проблеми: нечіткі жанрові межі, складність в інтерпретації, обмеження CNN та RNN.
- Необхідність у моделях, що поєднують точність, швидкодію та пояснюваність.

Мета

- Розробка гібридної моделі глибокого навчання для класифікації музичних жанрів.
- Інтеграція NMF (виділення джерел), CNN (виділення патернів), Temporal Attention (фокус на ключових моментах).



3

ОГЛЯД ІСНУЮЧИХ РІШЕНЬ

Методологія	Точність
CNN	57.5–87%
CNN з MFCC	76%
CNN + RNN (LSTM/GRU)	до 89.3%
CNN + DNN	90–92%
KNN / SVM / RF (класичні ML моделі)	66–76%
KNN (з MFCC або покращеними ознаками)	до 92.7%
Bottom-Up Broadcast CNN (BBNN)	93.9%
Запропонована модель (CNN + NMF + Attention)	93%

4

ОЗНАКИ АУДИО ДЛЯ КЛАСИФІКАЦІЇ

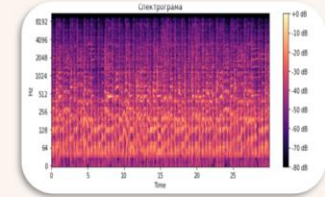
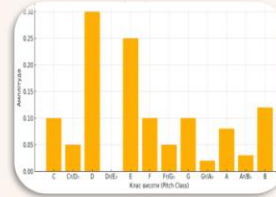
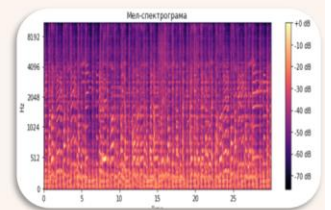
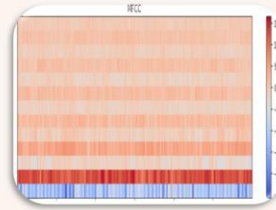
Для класифікації аудіожанрів використовуються різноманітні ознаки, що дозволяють описати звуковий сигнал з різних аспектів. Серед них:

ЧАСОВІ ТА ЧАСТОТНІ ОЗНАКИ

- **MFCC (МЕЛ-ЧАСТОТНІ КЕПСТРАЛЬНІ КОЕФІЦІЄНТИ):** Широко використовуються для розпізнавання мови та музики, відображаючи амплітудну огинаючу спектру.
- **СПЕКТРОГРАМИ ТА МЕЛ-СПЕКТРОГРАМИ:** Візуалізують зміни частотного складу звуку в часі, що є ключовим для ідентифікації тембру та текстури.

ТОНАЛЬНІ ХАРАКТЕРИСТИКИ

- **CHROMA (ХРОМА-ВЕКТОРИ):** Відображають інтенсивність кожної з 12 хроматичних висот (нот) у музичному фрагменті, незалежно від октави, що робить їх корисними для аналізу гармонії.
- **TONNETZ:** Геометрична модель для візуалізації зв'язків між тонами та акордами, допомагає аналізувати тональну структуру музики.



5

РЕАЛІЗАЦІЯ ТА ІНСТРУМЕНТИ



Python

Основна мова програмування.



Jupyter Notebook

Інтерактивне середовище для досліджень, коду і візуалізації.



Librosa

Витяг ознак (MFCC, спектрограми, chroma, тональний центроїд).



Pandas

Обробка таблиць ознак (CSV), робота з DataFrame.



NumPy

Числові обчислення та масиви.



Scikit-learn

Базові моделі (KNN, Naive Bayes, Decision Tree).



TensorFlow + Keras

Створення та навчання нейронних мереж (CNN, attention).



Matplotlib, Seaborn

Побудова графіків, теплових карт, спектрограм.

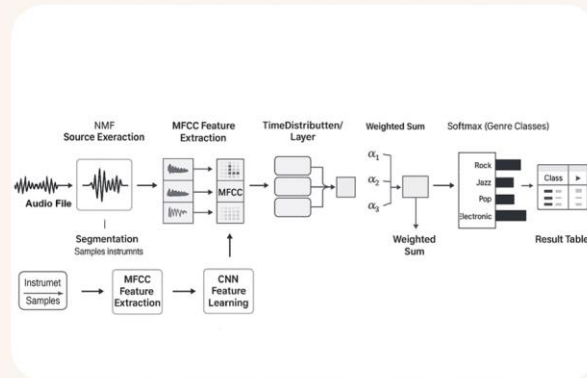


IPython.display

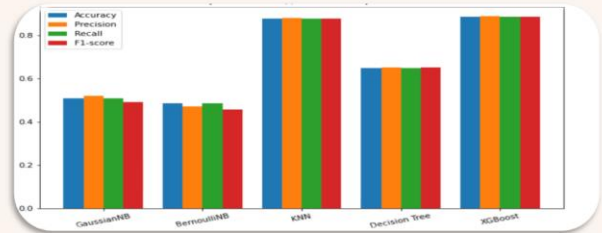
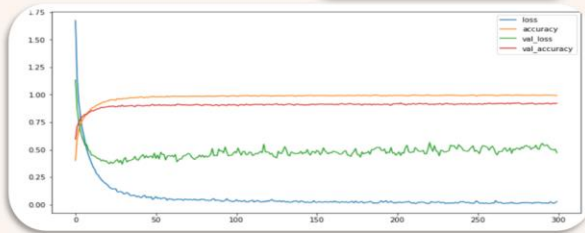
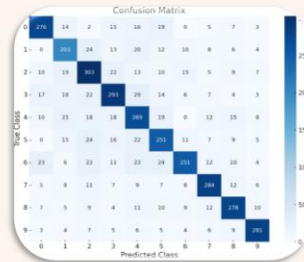
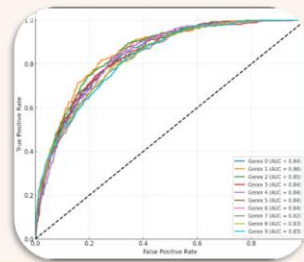
Інтеграція аудіо-прослуховування у ноутбучі.

6

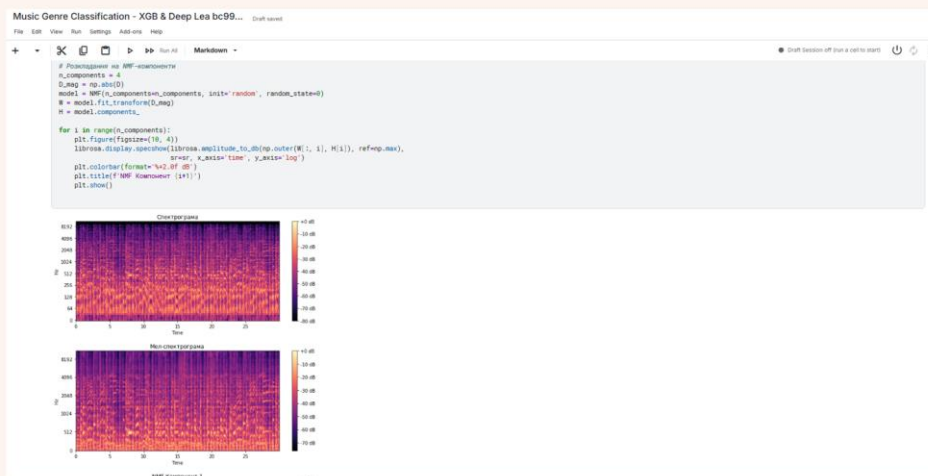
АРХІТЕКТУРА МОДЕЛІ



ПРОВЕДЕННЯ ТЕСТУВАНЬ



ДЕМОНСТРАТИВНЕ ВІДЕО



9

НАПРЯМКИ ПОКРАЩЕННЯ

Розширення тестування

Проведення тестування на додаткових датасетах, таких як FMA або MagnaTagATune, для підвищення надійності моделі.

Розробка методів аугментації

Створення та впровадження ефективних методів аугментації для підвищення стійкості до шумів та варіацій у музиці.

Вивчення Transformers

Дослідження можливостей використання архітектур Transformers для моделювання довготривалих залежностей у музичних треках

Інтеграція Explainable AI

Впровадження методів Explainable AI (XAI) для покращення інтерпретованості та підвищення довіри до результатів моделі.

10

ВИСНОВКИ

Гібридна модель

Побудовано CNN + NMF + Temporal Attention модель для класифікації аудіо.

Висока точність

Досягнуто **93% точності**, що є співставним з провідними сучасними підходами.

Інтерпретованість

Забезпечено **інтерпретованість** результатів за допомогою NMF-компонентів та механізму уваги.

Ефективність на GTZAN

Показана **ефективність моделі на реальному датасеті GTZAN**.

Комплексний цикл розробки

Розроблено повний цикл: від виділення ознак до оцінки результатів моделі.

11

АПРОБАЦІЯ ОТРИМАНИХ РЕЗУЛЬТАТІВ



ТЕЗИ ДОПОВІДІ

1. Жук М.В., Сергородцев І.Д., Фесенко Т.Г. Концепт архітектури web-застосунку для пошуку та прослуховування музичних композицій / Радиоелектроніка та молодь у ХХІ столітті: матеріали 28-го Міжнар. молодіж. форуму, 16–18 квітня 2024 р. –Харків: ХНУРЕ, 2024. –Т. 5. –С. 19–20.
2. Сергородцев І.Д., Жук М.В., Фесенко Т.Г. Особливості реалізації спільної фільтрації в Music Recommendation System / Радиоелектроніка та молодь у ХХІ столітті: матеріали 28-го Міжнар. молодіж. форуму, 16–18 квітня 2024 р. –Харків: ХНУРЕ, 2024. –Т. 5. –С. 63–65.