

# Aplikasi Rekomendasi Musik Berdasarkan Klasifikasi Genre Menggunakan Convolutional Neural Networks

Dewa Sukietra

Informatika, Fakultas Ilmu Komputer dan Desain, Universitas Kalbis  
Jalan Pulomas Selatan Kav.22, Jakarta 13210  
Email: theodorusdewa.s@gmail.com

**Abstract:** Convolutional Neural Network can be implemented on various deep learning approach for scientific or problem solving real world case. This research discusses implementing Convolutional Neural Network for music genre classification. Researcher will classified 8 genres music on this research. Music will be converted first to MFCC so its features can be extracted. Researcher use GTZAN Dataset as the base line for Convolutional Neural Network to learn each feature on a bunch of genres. Validation Accuracy with GTZAN dataset is 79% and Real world classification resulting in 60%. The classification program can be used on android devices. This research also serve music suggestions based the result of predicted genre.

**Keywords:** Music, Genre, CNN, classification, React Native, librosa, spotify, GTZAN

**Abstrak:** Convolutional Neural Network dapat di implementasikan dalam beragam metode deep learning baik untuk kajian ilmiah maupun pemecahan suatu masalah yang terjadi di dunia nya. Penelitian ini akan membahas bagaimana mengimplimentasikan Convolutional Neural Network dalam mengklasifikasikan genre musik. Peneliti akan mengklasifikasikan 8 buah genre pada penelitian ini. Musik akan terlebih dahulu di konversi menjadi MFCC agar fitur – fitur yang terdapat dalam musik dapat di ekstraksi. Peneliti menggunakan GTZAN dataset untuk menjadi tolak ukur dalam mempelajari fitur – fitur yang terdapat dalam masing – masing genre. Di dapatkan akurasi validasi sebesar 79% dan akurasi dengan data baru sebesar 60%. Hasil dari penelitian ini dapat di gunakan dengan menggunakan perangkat android. Penelitian ini juga akan memberikan rekomendasi musik berdasarkan hasil dari prediksi genre yang di lakukan.

**Kata kunci:** Musik, Genre, CNN, klasifikasi, React Native, librosa, spotify

## I. PENDAHULUAN

Musik merupakan suatu media yang di gunakan oleh manusia dalam mengungkapkan suatu pesan, ekspresi ataupun perasaan dari suatu individu manusia. Pada awalnya, musik hanya di kenal sebagai suatu kesenian yang menyusun nada – nada atau suara dari instrumen – instrumen yang ada sehingga menghasilkan suatu irama dan keharmonisan. Apapun yang menghasilkan bunyi atau suara dapat di anggap musik.

Namun seiring perkembangan zaman

dan kemajuan peradaban manusia, manusia kian mengelompokkan musik – musik yang memiliki kemiripan melodi antar satu sama lain ke dalam suatu aliran musik atau yang biasa kita kenal dengan nama genre musik. Maka dari itu, pada saat ini kita mengenal beragam macam jenis genre music. menggunakan huruf kecil. Penomoran tersebut tidak mnggunakan auto numbering.

Genre musik dapat di definisikan dari instrumen yang di gunakan, teknik musik, pola musik, tema dan beragam hal lainnya. Musik bagi sebagian individu dapat menjadi media refleksi yang dapat

meningkatkan konsentrasi, mengurangi depresi dan menjaga kesehatan mental. Memahami preferensi genre musik dapat membantu suatu individu manusia dapat lebih menikmati listening session dirinya dan dapat membuka wawasan yang lebih luas terhadap keberagaman musik – musik yang tersedia.

Di era modern sekarang terdapat beragam genre musik yang umum di temukan. Di tengah bervariasi nya jenis musik yang ada, penikmat musik pemula atau yang baru ingin memulai mengeksplor lebih dalam musik terkadang mengalami kesulitan dalam menentukan apa preferensi individu tersebut. Adakalanya suatu individu mengetahui satu musik yang pernah di dengar nya dan memiliki keterikatan atau ketertarikan terhadap musik tersebut namun tidak mengetahui apa jenis musik tersebut. Musik dengan karakter yang mirip dengan musik yang di gemari oleh individu tersebut tentu dapat menjadi jendela baru dalam mendalami bidang hobi ini. Namun di karenakan keterbatasan pengetahuan dan pengalaman hal ini menjadi penghalang dalam meningkatkan pengalaman yang menyenangkan dalam mendengarkan musik.

Penelitian ini berfokuskan untuk melakukan pembelajaran pada komputer dalam mengenali beragam genre musik sehingga dapat memberikan informasi tersebut ke pada pengguna dan memberikan rekomendasi musik yang memiliki kemiripan agar memudahkan penikmat musik pemula dalam menemukan preferensi musik yang sesuai dengan dirinya.

Dengan permasalahan yang ada, peneliti memiliki gagasan untuk membuat suatu aplikasi mobile yang dapat mendeteksi genre dari musik – musik yang di inputkan oleh pengguna. Aplikasi ini akan dapat membedakan masing – masing dari genre musik yang umum, dan akan memberikan rekomendasi dengan genre yang sama dengan harapan pengguna akan menyukai

rekomendasi tersebut dan menambah wawasan pengguna dalam dunia permusikan.

Dalam penelitian ini, akan di lakukan pendeteksian genre – genre yang umum seperti jazz, pop, rock, metal, classical, country, dan blues. Hasil rekomendasi akan menggunakan Spotify api untuk mendapatkan hasil rekomendasi berdasarkan hasil deteksi yang di dapatkan. Penelitian terdahulu dengan judul Music Genre Classification Using Convolutional Neural Network Akurasi yang di dapatkan dengan menggunakan algoritma Convolutional Neural Network dengan data set Tzanetakis1 menghasilkan akurasi sebesar 72.4%[1].

## **II. METODE PENELITIAN**

### **A. Tinjauan Pustaka**

#### **1. Kecerdasan Buatan**

Kecerdasan buatan merupakan suatu program automaton yang memiliki kemampuan untuk belajar dan mampu melaksanakan pekerjaan – pekerjaan yang membutuhkan kecerdasan manusia.[6]. Dari pengertian di atas, dapat di mengerti bahwa kecerdasan buatan itu sendiri di ciptakan untuk memberikan kemampuan ke pada mesin atau komputer agar dapat memiliki kemampuan berpikir, membuat keputusan, dan bertindak layaknya seperti manusia. Kecerdasan buatan memungkinkan mesin untuk membantu atau menggantikan peranan manusia dalam menyelesaikan suatu permasalahan atau pekerjaan.

Kecerdasan buatan sendiri adalah suatu istilah yang menyangkut pada otomatisasi suatu pekerjaan dengan mesin, dan kecerdasan buatan itu sendiri memiliki banyak kategori atau ilmu – ilmu yang di gunakan pada masing – masing bidang pekerjaan atau permasalahan.

## 2. Machine Learning

Machine Learning merupakan sebuah sub ilmu dari kecerdasan buatan di mana, cabang ilmu ini berfokus pada bagaimana membuat suatu mesin dapat memahami suatu pengetahuan dari hasil observasi yang di lakukan secara ulang – berulang. Machine learning dapat memahami suatu pola ataupun ilmu dengan cara melakukan pelatihan menggunakan suatu kumpulan data yang sudah di sediakan. Dari pola yang di dapatkan dari kumpulan data tersebut, mesin di harapkan dapat membuat suatu analisa dan menghasilkan suatu keluaran.

Untuk melatih suatu mesin, terdapat 3 kategori atau metode dalam melakukan pelatihan tersebut yaitu supervised learning, unsupervised learning, dan reinforcement learning[2]. Metode supervised learning adalah metode klasifikasi yang menandai suatu label data agar dapat di klasifikasikan. Sedangkan metode unsupervised learning tidak menandai suatu data dengan label tertentu, tapi mesin sendirilah yang akan mengelompokkan suatu kumpulan data yang memiliki kesamaan atau hubungan. Reinforcement learning merupakan metode yang menggabungkan kedua metode sebelumnya.

## 3. Convolutional Neural Networks

Convolutional Neural Network merupakan suatu cabang ilmu dari machine learning di mana CNN bekerja dengan cara melakukan operasi konvolusi yang menggabungkan beberapa lapisan pemrosesan dan berusaha untuk menyerupai jaringan syaraf biologis dalam menghasilkan suatu keluaran[3].

CNN memiliki beberapa lapisan – lapisan atau layer yaitu :

### a. Convolutional Layer

Merupakan lapisan yang memiliki parameter – parameter tertentu dari sebuah filters atau kernel yang akan menerima sebuah input, di mana setiap input yang melewati filter akan menciptakan suatu pola dan kumpulan pola – pola tersebut akan menjadi keluaran dari convolutional layer.

### b. Non-linearity Layer

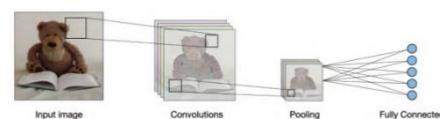
Mengaplikasikan beberapa fungsi atau feature map yang di buat oleh convolutional layer untuk membuat fungsi yang bersifat non-linear. Rectified Linear Unit(ReLU) merupakan salah satu cara pengaplikasian pada lapisan ini.

### c. Pooling Layer

Lapisan yang mengurangi dimensi dari feautre map yang di hasilkan sehingga dapat mempersingkat waktu komputasi yang di butuhkan. Lapisan ini juga bertugas untuk melakukan proses overfitting

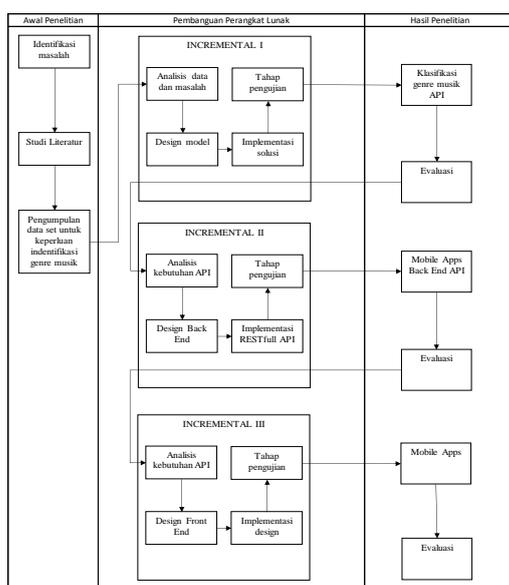
### d. Fully Connected Layer

Menghubungkan semua neurons yang ada pada lapisan – lapisan sebelumnya ke dalam satu lapisan sehingga membentuk suatu jaringan.



Gambar 1 Ilustrasi Fully Connected Layer

## B. Metodologi Penelitian



Gambar 2 Alur Kerja Penelitian

Kerangka pikiran diatas menggambarkan alur kerja perancangan aplikasi yang akan di buat. Tahap inkremental pertama merupakan tahapan untuk melatih jaringan syaraf tiruan dengan menggunakan dataset GTZAN. Datasets ini umum di gunakan untuk melakukan pelatihan model dalam pendeteksian atau pemrosesan musik[4]. Dataset ini terdiri dari 10 genre musik di mana setiap kelompok genre tersebut memiliki data audio sebanyak 100 yang memiliki durasi waktu 30 detik[5]. Hasil dari tahapan ini adalah model dari proses pembelajaran mesin tersebut dan API untuk mengakses model yang sudah di ekstrasi sehingga dapat di gunakan oleh aplikasi perangkat lunak lainnya.

Pada tahap inkremental kedua, peneliti akan membuat sebuah microservices sebagai basis backend dari aplikasi mobile yang akan di buat nantinya. Microservices ini terdiri dari fungsi fungsi untuk mengakses API dari hasil pembelajaran model sebelumnya, jalur pertukaran data antara mobile apps dan API model klasifikasi musik, dan beberapa third parties API untuk menampilkan rekomendasi musik dari hasil klasifikasi yang di dapatkan. Penggunaan third parties

API untuk memberikan rekomendasi musik berdasarkan genre yang di dapatkan dari klasifikasi adalah untuk memudahkan proses development di karenakan waktu yang terbatas. Microservices ini akan di buat mengikuti kaidah – kaidah RESTfull API yang baik untuk mendukung sebagai back-end dari mobile apps yang akan di buat pada inkremental ketiga. Untuk mendukung proses development pada tahapan ini, penulis akan menggunakan frameworks Express.js sebagai basis kode untuk back-end. Express.js.

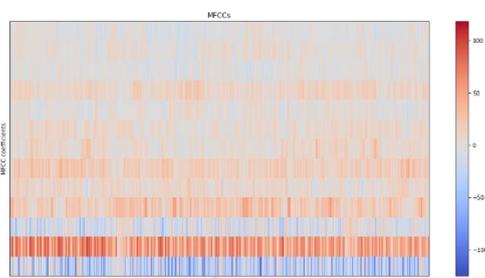
Tahapan inkremental ketiga adalah tahapan terakhir dari pengembangan sistem aplikasi dari penelitian ini. Pada tahapan ini, penulis akan merancang tampilan antarmuka berbasis mobile apps. Perancangan ini dimaksudkan untuk mempermudah pengguna dalam menggunakan aplikasi ini agar dapat di gunakan sehari – hari yang menjadi tujuan dari penelitian ini. Hasil perancangan antar muka yang sudah di buat kemudian akan di integrasikan dengan API – API yang sudah tersedia dari back-end services yang sudah di buat pada tahapan inkremental kedua sebelumnya. Pada tahapan ini, peneliti menggunakan frameworks dari javascript untuk aplikasi mobile yaitu react native. Penggunaan react native didasari oleh penggunaan fullstack javascript pada aplikasi ini agar mempermudah proses development dan mengurangi banyaknya bahasa pemrograman yang di gunakan. Hal ini merupakan salah satu keuntungan javascript dari bahasa pemrograman lainnya, yaitu dapat di gunakan secara fullstack untuk hal yang sederhana sampai dengan kompleks dalam membangun suatu sistem aplikasi.

### III. HASIL DAN PEMBAHASAN

#### A. Inkremental Pertama

Pertama – tama, langkah pertama yang

akan di lakukan adalah preprocessing data. Pada langkah ini, akan di lakukan pengestraksikan fitur pada setiap musik yang terdapat pada data set dengan cara mengkonversikan sinyal audio musik tersebut ke dalam Mel-frequency cepstral coefficients(MFCCs). Tahapan ini sangat penting untuk di lakukan agar di dapatkan sebuah data berupa gambar yang dapat di mengerti oleh arsitektur CNN yang akan di buat nantinya. Dalam proses tersebut, peneliti menggunakan library librosa. Sebagai ilustrasi, berikut salah satu sampel MFCCs dari dataset yang di gunakan.



Gambar 3 Visualisasi MFCCs

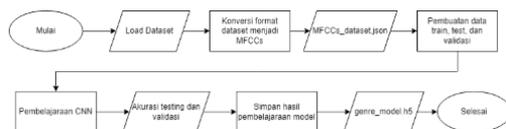
Tahapan selanjut nya adalah menyimpan MFCCs ke dalam suatu format yang dapat mudah di akses ketika di lakukan proses training. Peneliti menyimpan MFCCs tersebut ke dalam format json, di mana di dalam json akan terdapat key – key yang mewakili masing – masing genre dan pair values dari masing – masing key tersebut. Dengan format seperti ini, dapat mereplikasi struktur dataset aslinya sehingga mudah di mengerti dan format json adalah format yang universal di gunakan dalam pertukaraan data antar aplikasi.

Kemudian, akan di lakukan pelatihan terhadap mesin dengan menggunakan CNN. MFCCs yang sebelum nya sudah di siapakan akan di gunakan untuk mengesktraksi fitur – fitur yang akan di gunakan untuk mengidentifikasi genre musik. Proses pelatihan di mulai dengan cara membaca data json yang berisi informasi mengenai MFCCs dari masing – masing genre, dan membuat logika

program agar dapat melakukan pelatihan secara otomatis berdasarkan data yang di miliki.

Untuk arsitektur CNN yang di gunakan, peneliti me desain 3 convolutional layer, 1 layer dense dan dropout untuk meminimalisir terjadinya overfitting dan output layer sebagai keluaran nya. Peneliti melakukan beragam konfigurasi parameter dan epoch dalam tahapan ini untuk menemukan konfigurasi yang paling baik dalam kasus kali ini.

Setelah tahap pelatihan selesai, maka hasil dari pembelajaran tersebut akan di validasi. Hal ini akan menciptakan nilai akurasi pada model yang sudah di buat. Peneliti juga melakukan plotting pada data epoch, akurasi, dan validasi akurasi untuk menganalisa kemampuan model yang di buat. Peneliti juga melakukan percobaan validasi tambahan, dengan cara mem prediksi secara random musik yang terdapat di dalam data set untuk melihat secara real akurasi yang di miliki model. Hasil pembelajaran akan di simpa daalam format .h5 agar di mungkinkan untuk merekontruksi hasil pembalajara nyang sudah di lakukan.

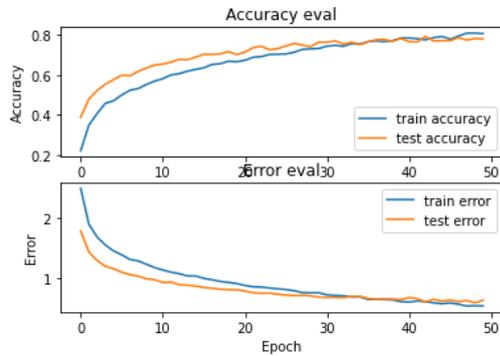


Gambar 4 Alur inkremental

Pada penelitian ini, penulis tidak memiliki referensi khusus yang sesuai dengan scenario sehingga penulis tidak memiliki nilai epoch dan batch yang optimal. Maka dari itu, peneliti mencoba beberapa nilai agar dapat mendapatkan nilai yang cocok untuk menghasilkan model yang optimal. Dengan data yang sedikit, penulis berusaha untuk meminimalisir jumlah epoch yang di lakukan agar terhindar dari overfitting.

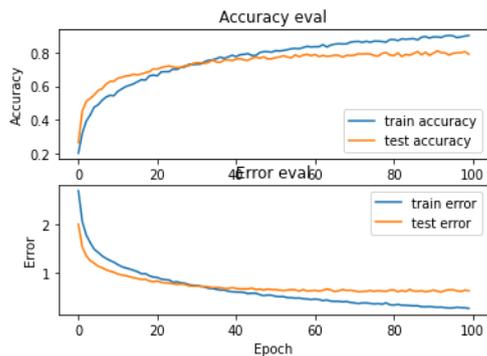
Pada percobaan pertama, nilai epoch

dan batch yang di gunakan adalah 50 dan 32 dengan grafik akurasi sebagai berikut :



Gambar 5 Akurasi Percobaan Pertama

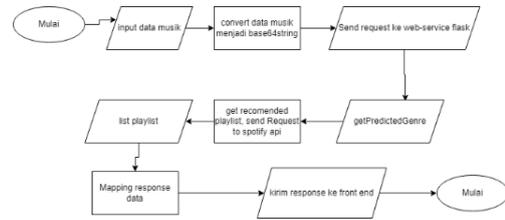
Pada percobaan kedua penulis menggunakan epochs sebesar 100, dengan grafik akurasi sebagai berikut :



Gambar 6 Akurasi Percobaan Kedua

### B. Inkremental Kedua

Inkremen kedua adalah tahapan perancangan untuk sisi server-side atau back-end dari aplikasi mobile yang akan di buat pada tahapan incremental ketiga nanti nya. Tahapan ini terdiri dari proses handling data request dan response dari aplikasi mobile dengan web-service python yang di buat sebelumnya, dan mendapatkan rekomendasi musik berdasarkan genre yang di deteksi dengan cara meng consume web API dari Spotify.



Gambar 7 Alur Kerja Inkremental 2

The screenshot shows a REST client interface with a POST request to the endpoint `/api/genre`. The response body is displayed in JSON format:

```

    {
      "message": "Operation success",
      "status": 200,
      "data": {
        "message": "Success",
        "predictedGenre": "pop",
        "recommendation": [
          {
            "title": "Young Money",
            "songName": "Trapster",
            "album": "Bliss of An Empire (Deluxe Edition)",
            "preview": "https://open.spotify.com/track/6H2L116x6d37F8zx7k",
            "albumCover": "https://i.scdn.co/image/ab67616d0000b30072863c37377899e9f7b4554"
          },
          {
            "title": "Twenty One Pilots",
            "songName": "Jumpcut",
            "album": "Irevel",
            "preview": "https://open.spotify.com/track/3F1uW86d05d3z26q7",
            "albumCover": "https://i.scdn.co/image/ab67616d0000b30072863c37377899e9f7b4554"
          },
          {
            "title": "Drake",
            "songName": "Hold On, We're Going Home"
          }
        ]
      }
    }
  
```

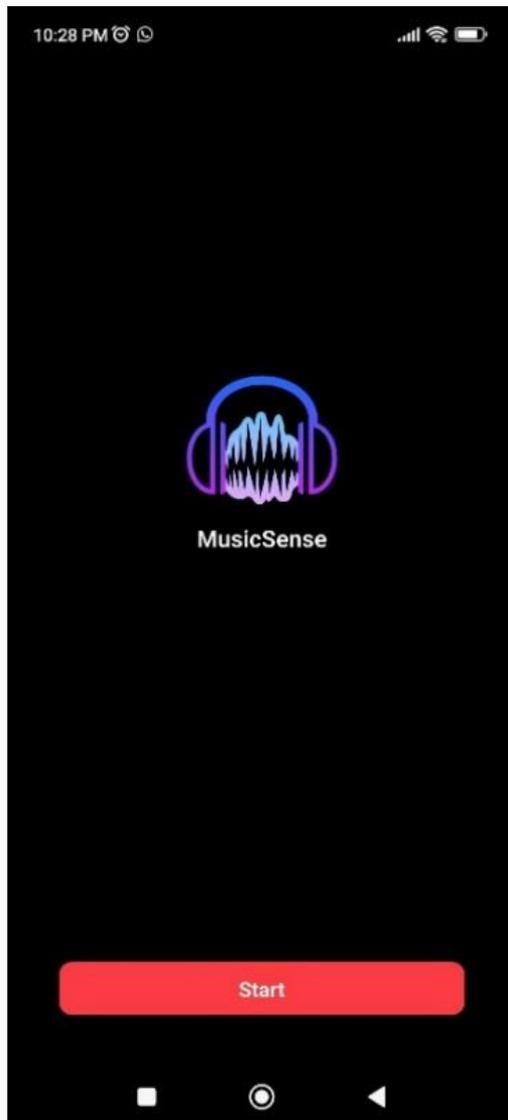
Gambar 8 Keluaran kode program inkremental dua

### C. Inkremental Ketiga

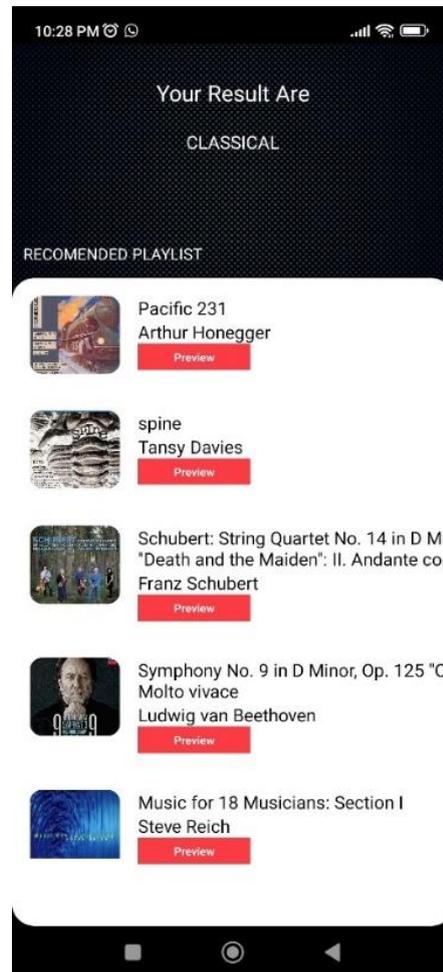
Proses analisis yang akan di lakukan pada tahap incremental tiga ini adalah mencoba menterjemahkan keluaran dari API yang di buat pada inkremen dua, agar dapat menjadi informasi yang mudah di cerna dan di gunakan oleh pengguna sehingga tahapan ini berfokus pada tampilan program.

Penulis akan mengevaluasi hasil keluaran dari program inkremen 2, memberikan sebuah antarmuka yang dapat menyajikan informasi yang sudah di dapatkan oleh program sebelumnya dan juga memikirkan kemudahan penggunaan aplikasi sehingga informasi yang di dapatkan oleh aplikasi GUI ini dapat di sampaikan dengan se jelas – jelasnya.

Dari analisa yang sudah di lakukan, penulis akan membuat tampilan antar muka yang memiliki komponen – komponen yang mudah di gunakan seperti tombol, dan sebuah card yang akan berisi muatan informasi yang di dapatkan dari API pada inkremen dua.



Gambar 9 Halaman Awal



Gambar 10 Halaman Hasil

Gambar 10 merupakan halaman antarmuka hasil deteksi genre musik yang di unggah oleh pengguna. Halaman ini terdiri dari genre dari musik yang di deteksi, beserta rekomendasi musik. Rekomendasi musik itu sendiri berdasarkan dari genre yang didapatkan oleh program, di mana penulis menyediakan detail dari musik – musik rekomendasi yang di dapatkan yaitu judul, singer, cover album dan link preview lagu tersebut.

#### IV. SIMPULAN

CNN dapat di gunakan untuk mengklasifikasi genre music dengan menggunakan data MFCC dari sebuah file audio. Untuk dapat melakukan pelatihan

mesin dengan CNN, di perlukan data gambar agar dapat di ekstraksi fitur – fitur. Maka dari itu data musik asal di konversi terlebih dahulu menjadi MFCC. Peningkatan epoch pada penelitian ini tidak meningkatkan akurasi validasi, di sebabkan dengan minim nya data pada setiap genre yang hanya 70 buah menimbulkan rentan nya overfitting.

Hasil percobaan dengan data baru memiliki nilai yang berbeda dari hasil test, hal ini di karenakan GTZAN dataset hanya memiliki musik – musik dengan ciri khas era 1990 – 2000. Seiring nya waktu, musik terus berkembang sehingga ada ketidak sesuaian persepsi terhadap suatu genre sehingga model kesulitan untuk mendeteksi dengan benar genre musik modern.

Hasil rekomendasi musik pada penelitian kali ini menggunakan web-service api spotify yang memiliki koleksi lagu yang massif.

## DAFTAR RUJUKAN

- [1] Q. Kong, “Music Genre Classification Using Convolutional Neural Network,” pp. 0–1, 2014.
- [2] A. Roihan, P. A. Sunarya, and A. S. Rafika, “Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper,” *IJCIT (Indonesian J. Comput. Inf. Technol.*, vol. 5, no. 1, pp. 75–82, 2020, doi: 10.31294/ijcit.v5i1.7951.
- [3] J. Brownlee, “A Gentle Introduction to Batch Normalization for Deep Neural Networks,” 2019.
- [4] B. L. Sturm, “The GTZAN dataset: Its contents, its faults, their effects on evaluation, and its future use,” no. 11, pp. 1–29, 2013, doi: 10.1080/09298215.2014.894533.
- [5] A. OLTEANU, “GTZAN Dataset - Music Genre Classification,” 2020. <https://www.kaggle.com/datasets/andradaoltEANU/gtzan-dataset-music-genre-classification>.