

LAPORAN TUGAS BESAR 2

IF2123 ALJABAR LINEAR DAN GEOMETRI



Kelompok 14 Menjajal Nilai Maksimal

Najwa Kahani Fatima 13523043

Mayla Yaffa Ludmilla 13523050

Nayla Zahira 13523079

SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA

INSTITUT TEKNOLOGI BANDUNG

2024

DAFTAR ISI

BAB I.....	4
I. Pemrosesan Suara dan Gambar.....	4
II. Information Retrieval.....	5
BAB II.....	6
BAB III.....	8
BAB IV.....	9
BAB V.....	10

BAB I

DESKRIPSI MASALAH

I. Pemrosesan Suara dan Gambar

Suara selalu menjadi hal yang paling penting dalam kehidupan manusia. Manusia berbicara mengeluarkan suara dan mendengarkan suatu suara untuk diresap ke otak dan mencari informasi dari suara tersebut. Suara juga bisa dijadikan orang-orang di dunia ini sebuah media untuk membuat karya seni. Contohnya adalah alat mendeteksi lagu. Manusia bisa mendeteksi suara dengan menggunakan indera pendengar dan memberikan kesimpulan akan apa jenis suara tersebut melalui respon dari otak. Sama seperti manusia, teknologi juga bisa mendeteksi suara dan memberikan jawaban mereka melalui algoritma-algoritma yang beragam bahkan bisa melebihi kapabilitas manusia. Dengan menggunakan algoritma apapun, konsep dari pendeteksi dan interpretasi suara itu bisa juga disebut dengan sistem temu balik suara atau bisa disebut juga dengan *audio retrieval system*. Banyak aplikasi yang menggunakan konsep sistem temu balik contohnya adalah Shazam.



Gambar 1.1 Shazam sebagai aplikasi *audio retrieval system*

Selain suara, manusia juga memiliki penglihatan sebagai salah satu inderanya dan bisa melihat warna dan gambar yang bermacam-macam. Teknologi komputasi juga memiliki

kapabilitas yang sama dan bisa melihat gambar sama seperti kita, tetapi teknologi seperti ini juga bisa merepresentasikan gambar tersebut sebagai beragam-ragam angka yang bisa disebut juga fitur. Tahun ke tahun, *image processing* selalu menjadi fokus utama dari tugas besar 2 Algeo. Algoritma yang digunakan adalah Eigenvalue, Cosine Similarity, Euclidean Distance, dll.

Anda sudah melewati Tugas Besar 1 yaitu tentang matriks dan implementasi terhadap berbagai hal. Matriks adalah salah satu komponen yang penting dalam aplikasi aljabar vektor. Di dalam Tugas Besar 2 ini, anda diminta untuk membuat semacam aplikasi Shazam yaitu sebuah aplikasi yang meminta input lagu dan aplikasi tersebut mendeteksi apa nama dari lagu tersebut dan beberapa detail lainnya. Pada tugas besar ini, anda akan menggunakan aljabar vektor untuk mencari perbandingan antar satu audio dengan audio yang lain. Anda akan menggunakan konsep yang bernama *Music Information Retrieval* atau MIR untuk mencari dan mengidentifikasi suara berdasarkan fitur-fitur yang dimilikinya. Tidak hanya itu, anda juga akan menggunakan konsep Principal Component Analysis (PCA) untuk mencari kumpulan audio melalui deteksi wajah berbagai orang (anggap saja mereka sebagai seorang penyanyi).

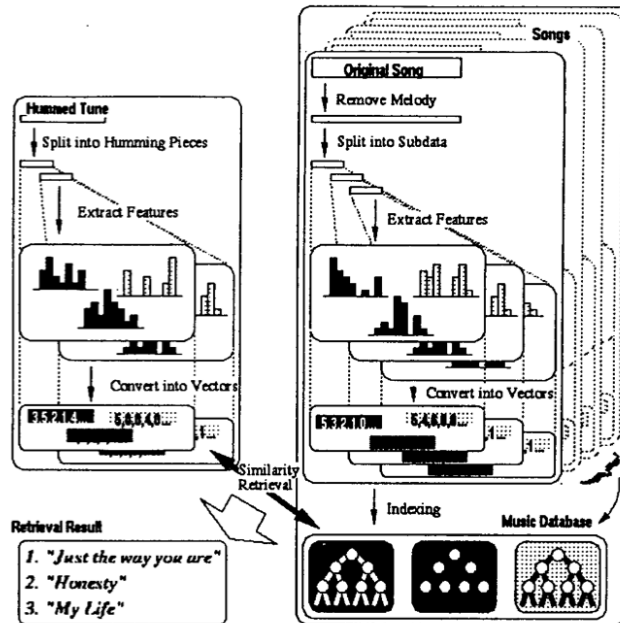
II. Information Retrieval

Information Retrieval adalah konsep meminta informasi dari sebuah data dengan memasukkan data tertentu. Pada tugas besar ini, anda akan berkitik dengan 2 jenis Information Retrieval. *Image Retrieval* dan *Music Information Retrieval*. *Image Retrieval* adalah konsep untuk memasukkan sebuah input gambar dan berharap mendapatkan gambar yang ada di data sesuai dengan informasi dan perhitungan yang diinginkan. Sedangkan *Music Information Retrieval* (MIR) adalah konsep untuk memasukkan sebuah input audio dan berharap mendapatkan audio yang ada di data sesuai dengan informasi dan perhitungan yang diinginkan. Pada tugas besar kali ini, kalian akan mengimplementasikan *Image Retrieval* dengan menggunakan Principal Component Analysis dan *Music Information Retrieval* dengan menggunakan humming.

BAB II

TEORI SINGKAT

I. Sistem Temu Balik Suara (MIR)



Music Information Retrieval (MIR) adalah konsep untuk mengambil informasi dari data audio menggunakan input audio tertentu. Dalam implementasinya, sistem MIR bertujuan untuk mencocokkan input audio, seperti humming atau rekaman suara, dengan audio yang tersedia di dalam basis data berdasarkan analisis fitur dan perhitungan kemiripan. Sistem ini relevan untuk berbagai aplikasi, seperti pencarian lagu, analisis musik, dan pengenalan melodi.

Dalam metode query by humming, pengguna cukup menyenandungkan melodi yang diingat, dan sistem akan mencocokkannya dengan melodi yang ada di basis data. Proses ini melibatkan beberapa langkah utama:

1. Pemrosesan Audio

Input berupa audio direkam atau diterima dalam bentuk file MIDI. Data ini diproses untuk mengekstrak informasi melodi utama, yang biasanya terdapat pada track di Channel 1. Teknik windowing digunakan untuk membagi melodi menjadi segmen-segmen kecil dengan ukuran 20–40 beat dan langkah pergeseran (sliding window) sebesar 4–8 beat.

Untuk meningkatkan akurasi pencocokan, dilakukan normalisasi pada tempo dan pitch. Normalisasi ini mengurangi pengaruh variasi dalam humming pengguna, seperti perbedaan kecepatan atau ketinggian nada. Notasi musik diubah menjadi representasi numerik yang mempertimbangkan durasi dan urutan nada, sehingga sistem dapat membandingkan potongan melodi dengan basis data.

Formula untuk normalisasi pitch:

$$NP(note) = \frac{(note - \mu)}{\sigma}$$

di mana μ adalah rata-rata pitch, dan σ adalah standar deviasi pitch.

2. Ekstraksi Fitur

Ekstraksi fitur adalah proses mengubah data audio menjadi representasi numerik berupa histogram yang mencerminkan karakteristik melodi. Sistem ini menggunakan tiga fitur utama:

1. Absolute Tone Based (ATB)

Fitur ini menghitung frekuensi kemunculan setiap nada pada skala MIDI (0–127) dan menghasilkan histogram dengan 128 bin. Histogram ini memberikan gambaran distribusi nada absolut, membantu menangkap karakteristik statis melodi. Prosesnya meliputi:

2. Relative Tone Based (RTB)

RTB fokus pada pola interval melodi dengan menganalisis perbedaan antara nada-nada yang berurutan. Fitur ini menghasilkan histogram dengan 255 bin pada rentang nilai -127 hingga +127. Pendekatan ini relevan untuk mencocokkan humming dengan melodi yang tidak bergantung pada nada absolut.

3. First Tone Based (FTB)

FTB menyoroti hubungan relatif antara nada dalam melodi dengan nada pertama sebagai referensi. Histogram yang dihasilkan mencerminkan struktur relatif nada, sehingga lebih stabil terhadap variasi pitch pengguna.

$$H_{norm} = \frac{H[d]}{\sum_d H[d]}$$

di mana H adalah histogram, dan d adalah bin dalam histogram.

3. Penghitungan Similaritas

Setelah fitur diekstraksi, langkah berikutnya adalah menghitung tingkat kemiripan antara melodi query dengan data dalam basis data. Setiap histogram diubah menjadi vektor, dan perhitungan similaritas dilakukan menggunakan cosine similarity.

Cosine Similarity mengukur kesamaan dua vektor dengan menghitung sudut antara keduanya. Semakin kecil sudutnya (nilai mendekati 1), semakin mirip kedua vektor. Formula cosine similarity:

$$\cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

II. Metode Ekstraksi Fitur Berdasarkan Humming

Ekstraksi fitur berdasarkan humming adalah aplikasi dari MIR (Music Information Retrieval) yang memungkinkan pengenalan musik meskipun hanya dengan menggunakan suara manusia yang berupa humming atau nyanyian tanpa lirik. Pendekatan ini memungkinkan seseorang untuk menyenandungkan melodi dan menemukan lagu yang sesuai dalam basis data musik. File humming akan diproses menggunakan metode windowing serta normalisasi tempo dan pitch untuk mengurangi variasi humming. Setelah itu, setiap note dikonversi menjadi representasi numerik yang mempertimbangkan durasi dan urutan nada sehingga sistem dapat membandingkan potongan melodi dengan dataset yang ada.

1. Pemrosesan audio

Pemrosesan audio dalam sistem *query by humming* menggunakan file MIDI dengan fokus pada track melodi utama, biasanya di Channel 1. Proses ini melibatkan metode windowing,

di mana melodi dibagi menjadi segmen sepanjang 20-40 beat dengan sliding window sebesar 4-8 beat. Teknik ini memungkinkan sistem untuk mencocokkan berbagai potongan lagu yang mungkin diingat pengguna.

Proses windowing juga mencakup normalisasi tempo dan pitch untuk mengurangi perbedaan dalam humming pengguna. Setiap *note event* dikonversi menjadi representasi numerik yang mempertimbangkan durasi dan urutan nada, sehingga memudahkan perbandingan potongan melodi dengan data di dalam basis data. Formula normalisasi tempo adalah.

$$NP(\text{note}) = \frac{\text{note} - \mu}{\sigma}$$

μ : Rata-rata pitch

σ : Standar deviasi pitch

2. Ekstraksi Fitur

- Absolute Tone Based (ATB)

ATB menghitung frekuensi kemunculan setiap nada dalam skala MIDI (0-127) dan menghasilkan histogram 128 bin. Histogram ini menunjukkan distribusi absolut nada dan membantu memahami karakteristik statis melodi.

- Relative Tone Based (RTB)

RTB menganalisis perubahan antar nada yang berurutan dan menghasilkan histogram dengan 255 bin dalam rentang -127 hingga +127. RTB berfokus pada pola interval melodi, yang lebih relevan untuk mencocokkan humming tanpa bergantung pada pitch absolut.

- First Tone Based (FTB)

FTB menganalisis perbedaan setiap nada dengan nada pertama, menciptakan histogram dengan 255 bin dalam rentang -127 hingga +127. Pendekatan ini menangkap hubungan relatif nada terhadap titik referensi awal, membantu mengatasi variasi pitch.

III. Image Retrieval dengan Principal Component Analysis

Image Retrieval adalah proses mencari gambar dalam basis data yang relevan dengan gambar query yang diberikan. Salah satu metode yang sering digunakan dalam pengolahan citra untuk mengurangi dimensi dan mengekstraksi informasi penting adalah Principal Component Analysis (PCA). PCA adalah teknik yang digunakan untuk mereduksi dimensi data dengan cara mengubah set data asli menjadi sekumpulan data baru yang mengandung informasi yang paling signifikan dari data asli. Dalam konteks image retrieval, PCA digunakan untuk mengurangi kompleksitas gambar dan mengekstrak fitur-fitur utama yang mencirikan gambar tersebut. Pada awalnya ekstraksi fitur citra dilakukan dengan cara mengubah gambar dataset menjadi array yang dapat dioperasikan. Setelah itu, data data yang paling signifikan dihasilkan dengan PCA. Query yang masuk akan dicocokkan dengan data hasil PCA dan dicari jarak euclidean terkecilnya untuk mencari data yang paling mirip dari dataset.

Untuk melakukan image retrieval menggunakan metode PCA, ada beberapa konsep yang diterapkan.

1. Singular Value Decomposition (SVD)

Singular Value Decomposition (SVD) adalah sebuah teknik dalam aljabar linier yang digunakan untuk memecah sebuah matriks menjadi tiga matriks yang lebih sederhana. Secara khusus, SVD menguraikan sebuah matriks A berukuran $m \times n$ menjadi tiga matriks:

$$A = U\Sigma V^T$$

- U adalah matriks ortogonal berukuran $m \times m$
- Σ adalah matriks diagonal berukuran $m \times n$ yang berisi singular value di sepanjang diagonalnya.
- V^T adalah transpose dari matriks ortogonal berukuran $n \times n$

2. Nilai Eigen dan Vektor Eigen

Eigenvalue (nilai eigen) adalah skalar yang menggambarkan seberapa banyak vektor tertentu diperbesar atau diperkecil ketika diterapkan pada matriks A. Eigenvector adalah vektor yang tidak berubah arah ketika matriks A diterapkan padanya. Eigenvalue mengukur perubahan skala pada vektor tersebut. Jika eigenvalue-nya besar, vektor eigen akan diperbesar secara signifikan, dan jika eigenvalue-nya kecil, vektor eigen akan diperkecil.

3. Matriks Kovarians

Matriks Kovarians adalah matriks yang menggambarkan hubungan antara dua atau lebih variabel dalam sebuah dataset. Setiap elemen dalam matriks kovarians menunjukkan kovarians antara sepasang variabel. Kovarians itu sendiri adalah ukuran sejauh mana dua variabel bergerak bersama; apakah keduanya meningkat atau menurun secara bersamaan, atau apakah ada hubungan negatif antara keduanya.

$$C = \frac{1}{n-1} X^T X$$

4. Data Centering

Data centering adalah proses mentransformasikan data sehingga rata-rata setiap fitur menjadi nol. Hal ini dilakukan dengan mengurangkan nilai rata-rata (mean) dari setiap fitur pada seluruh nilai fitur tersebut.

$$\mu_j = \frac{1}{N} \sum_{i=1}^N x_{ij}$$

BAB III

ARSITEKTUR WEBSITE

I. Arsitektur Website

Website MuseIc terdiri atas beberapa halaman, yaitu 1) halaman utama, 2) halaman pencarian berdasarkan musik, dan 3) halaman pencarian berdasarkan gambar. Pada halaman utama terdapat deskripsi singkat mengenai website dan fitur-fitur yang dibuat. Ditampilkan juga data kontributor dalam tugas besar ini. Navigation bar berisi opsi terkait fitur yang dapat digunakan, yaitu Music Retrieval dan Album Finder.

Pada fitur Music Retrieval dan Album Finder, terdapat beberapa bagian pada layar, yaitu sidebar dan komponen utama. Pada sidebar, terdapat komponen Upload Dataset Image, Upload Dataset Audio, Upload Mapper, tombol Load Dataset, tombol start, tampilan waktu yang dibutuhkan untuk memproses dataset, nama file yang telah di-load, pagination controller, serta banyaknya file yang ditampilkan. Pengguna akan mengunggah dataset gambar, audio, serta mapper dengan menekan tombol yang ada. Setelah dataset selesai di-load, pengguna dapat mengunggah query berdasarkan fitur yang digunakan. Ketika tombol start ditekan, program akan menjalankan model pencarian similaritas yang telah dibuat pada folder backend.

Pada saat Load Data, komponen utama akan menampilkan seluruh data pada dataset beserta banyaknya data yang ditampilkan. Pengguna dapat menekan tombol Next dan Previous untuk navigasi ke halaman selanjutnya. Ketika tombol start query ditekan, komponen utama akan menampilkan data dengan kemiripan minimal 70% relatif terhadap seluruh dataset.

II. Arsitektur Program Information Retrieval (backend)

Sehingga, pengembangan program ini dibagi menjadi 2 model utama, yaitu *music information retrieval model* (MIR model) dan *image information retrieval model* (IIR Model).

1. MIR Model

Pada bagian implementasi konsep sistem temu balik berbasis suara, terdapat beberapa fungsi yang dibuat untuk mendukung dalam proses mengekstrak dan mengolah informasi musikal. Berikut beberapa fungsi tersebut dan penjelasannya.

1. `extract_notes(midi_file_path)`

Fungsi ini bertujuan untuk mengekstraksi nada-nada (notes) dari file MIDI yang diberikan. Dalam implementasinya, fungsi ini membaca file MIDI menggunakan library `mido` dan memindai setiap track di dalamnya untuk mencari pesan dengan tipe `note_on` yang terdapat pada channel 0 (yang merepresentasikan channel 1 pada MIDI). Nada yang diambil adalah yang memiliki nilai `velocity` lebih dari 0, menandakan bahwa nada tersebut sedang dimainkan. Hasil akhir dari fungsi ini adalah daftar nada yang ditemukan pada file MIDI yang dapat digunakan untuk analisis lebih lanjut.

2. `normalize_pitch(notes)`

Fungsi ini digunakan untuk menormalisasi nada-nada yang telah diekstraksi agar siap untuk analisis statistik. Proses normalisasi dilakukan dengan menghitung rata-rata (`avg_pitch`) dan standar deviasi (`std_pitch`) dari daftar nada lalu diproses sesuai rumus berikut.

$$NP(note) = \frac{(note - \mu)}{\sigma}$$

3. `windowing(notes, window_size, step_size)`

Fungsi ini berfungsi untuk membagi daftar nada yang telah diekstraksi menjadi segmen-segmen kecil dengan panjang tertentu, yang disebut "jendela" (window). Proses ini membagi melodi menjadi segmen 20 beat dengan sliding window 4 beat.

4. `absolute_tone_based(arrMidi)`

Fungsi ini bertujuan untuk menghitung histogram distribusi nada absolut (absolute tone) dalam segmen nada tertentu. Histogram ini membagi rentang nilai

nada MIDI (0–127) menjadi 128 bin, lalu menghitung frekuensi kemunculan nada-nada dalam segmen tersebut.

5. `relative_tone_based(arrMidi)`

Fungsi ini berfungsi untuk menghitung histogram distribusi interval nada relatif (relative tone) antara nada-nada yang berurutan dalam segmen. Fungsi ini menghitung selisih (interval) antara setiap pasangan nada berurutan dan membagi rentang interval nada (-127 hingga 127) menjadi 255 bin

6. `first_tone_based(arrMidi)`

Fungsi ini menghitung histogram distribusi interval nada relatif terhadap nada pertama dalam segmen (first tone). Fungsi ini bekerja dengan menghitung selisih antara setiap nada dalam segmen dengan nada pertama, lalu membagi rentang interval (-127 hingga 127) menjadi 255 bin.

2. IIR Model

Pada bagian implementasi konsep sistem temu balik berbasis gambar, terdapat beberapa fungsi yang dibuat untuk mendukung dalam proses mengekstrak dan mengolah informasi musikal. Berikut beberapa fungsi tersebut dan penjelasannya.

1. `grayscale(filename)`

Fungsi ini bertugas untuk mengubah citra berwarna menjadi citra dalam skala abu-abu (grayscale), yang kemudian diubah ukurannya menjadi 144x144 piksel. Setelah itu, citra diubah menjadi vektor satu dimensi (flatten) sehingga dapat diproses lebih lanjut.

2. `standarize(image_arrays)`

Fungsi `standardize(image_arrays)` digunakan untuk menstandarisasi data citra dengan menghitung rata-rata (mean) dan standar deviasi (std) dari intensitas piksel di seluruh citra dalam basis data. Setiap piksel kemudian dinormalisasi menggunakan rumus:

$$standardized_value = \frac{pixel_value - mean}{std}$$

3. `comp_covariance(standardized_data)`

Fungsi ini berfungsi untuk menghitung matriks kovarians dari data citra yang telah distandarisasi. Matriks kovarians ini menggambarkan hubungan antar dimensi piksel dalam data, yang akan digunakan dalam analisis PCA untuk mengekstraksi komponen utama.

4. `comp_svd(covariance_data, k)`

Fungsi ini menggunakan Singular Value Decomposition (SVD) untuk memecah matriks kovarians menjadi tiga komponen utama: matriks singular kiri (U_k), nilai singular (S), dan matriks singular kanan (V^T). Parameter k menentukan jumlah komponen utama yang akan digunakan. Matriks U_k digunakan untuk mereduksi dimensi data.

5. `projection_data(standardized_data, U)`

Fungsi ini menghitung proyeksi data citra ke ruang komponen utama menggunakan matriks U_k yang diperoleh dari SVD. Hasil proyeksi ini berupa matriks yang merangkum informasi penting dari data citra dalam dimensi yang lebih rendah.

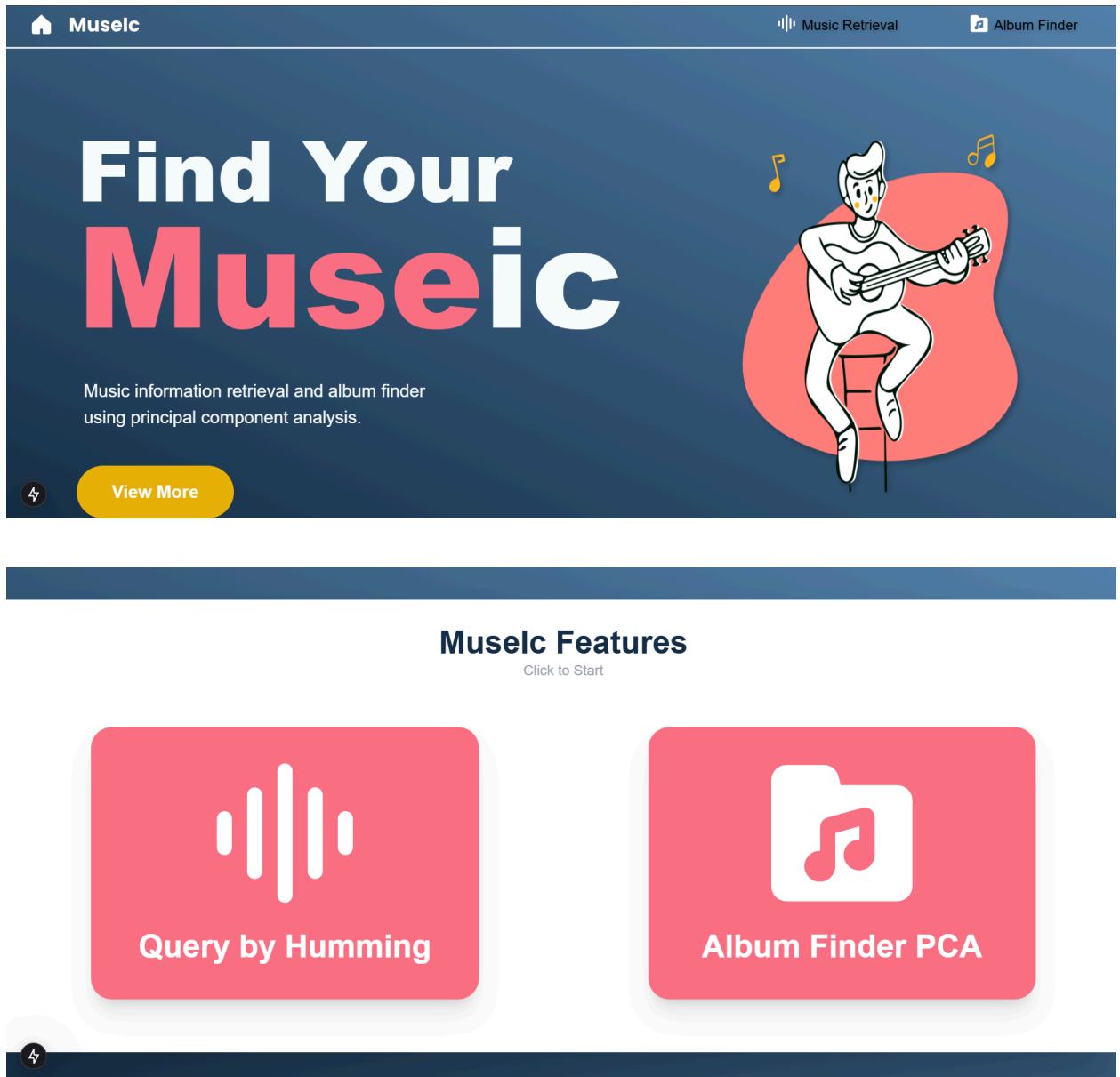
6. `euc_dist(projected_query, projected_data)`

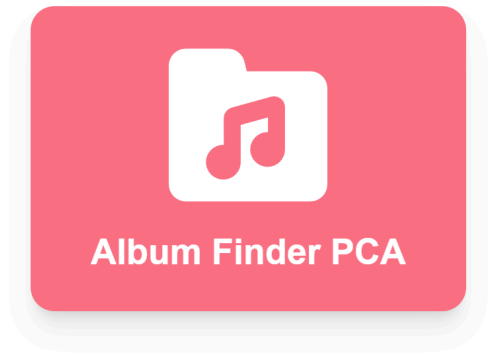
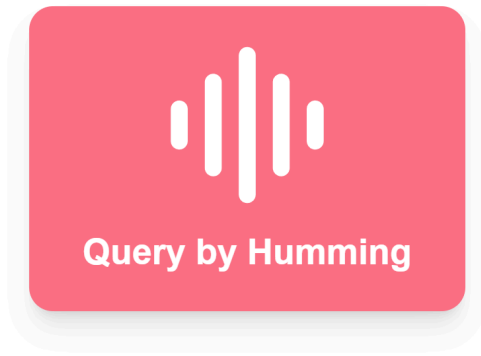
Fungsi `comp_covariance(standardized_data)` berfungsi untuk menghitung matriks kovarians dari data citra yang telah distandarisasi.

BAB IV

EKSPERIMEN

I. Landing Page





The Contributors

Kelompok Menjajal Nilai Maksimal

Najwa Kahani Fatima - 13523043 - K01

Mayla Yaffa Ludmilla - 13523050 - K02

Nayla Zahira - 13523079 - K01

II. Laman Music Retrieval

Muselc

Music RetrievalAlbum Finder

Upload Humming

Query: none

START

Dataset Audio

Selected: none

Dataset Image

Selected: none

Mapper

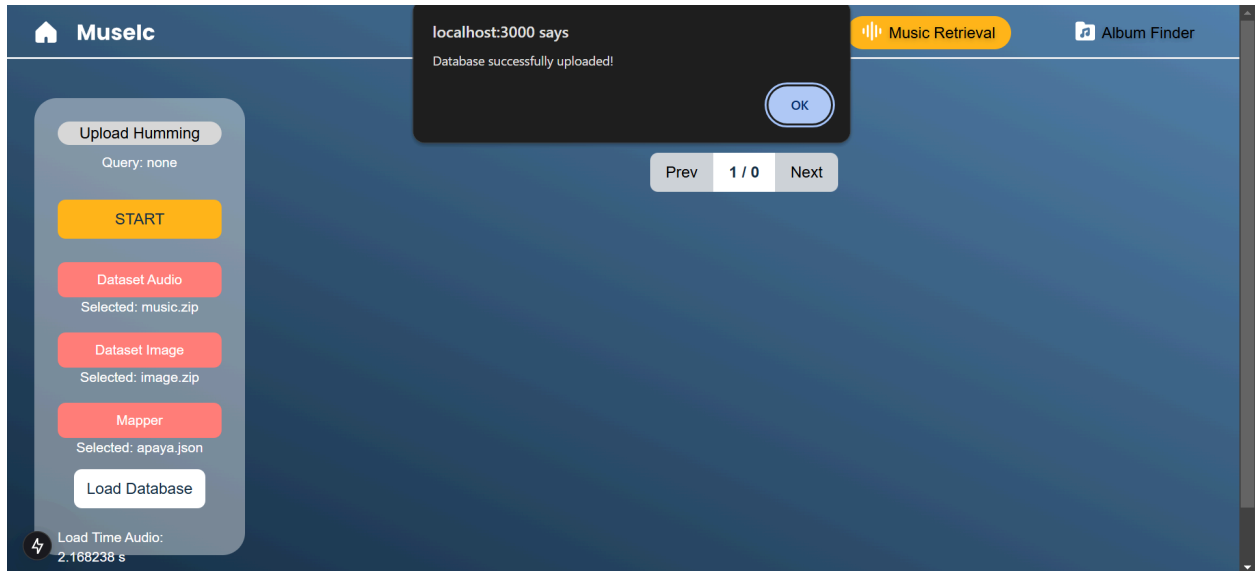
Selected: none

Load Database

Result: 0 Data

Prev1 / 0Next

Load Time Audio: Unknown



BAB V

KESIMPULAN

Berdasarkan materi yang telah dipelajari dalam kuliah IF2123 Aljabar Linear dan Geometri, kami mengimplementasikan konsep-konsep terkait pemrosesan vektor dan matriks dalam sebuah program mengenai Principal Component Analysis (PCA) dan Music Information Retrieval (MIR). Program yang kami buat dapat menyelesaikan berbagai permasalahan, seperti pencocokan gambar menggunakan PCA serta penerapan metode MIR untuk pencarian musik berdasarkan humming atau cuplikan audio. Melalui tugas ini, kami semakin memahami materi mengenai aljabar linear yang dibutuhkan dalam pembuatan program seperti SVD (Single Value Decomposition), euclidean distance, dan cosine similarity. Selain itu, kami juga belajar bagaimana cara mengintegrasikan dan mengimplementasikan teori dalam bentuk website dengan menggunakan bahasa pemrograman Python dan Javascript.

Selama pengerjaan tugas besar ini, kami menghadapi berbagai kendala, terutama terkait waktu eksekusi yang cukup lama dan tantangan dalam pembuatan website. Oleh karena itu, kami mengajukan beberapa saran untuk meningkatkan kualitas tugas ini ke depannya, yaitu:

1. Mengoptimalkan algoritma dengan menggunakan bahasa pemrograman lain yang memungkinkan eksekusi lebih cepat.
2. Mempertimbangkan penggunaan framework Front-End yang lebih ringan dan efisien untuk meningkatkan performa dan pengalaman pengguna.

LAMPIRAN

Tautan repository: <https://github.com/najwakahanifatima/Algeo02-23043>

Tautan video:

https://drive.google.com/file/d/1R7XpB3o3_pcjndssHeqU01POSBAC9rx7/view?usp=drive_sdk

