

LAPORAN TUGAS BESAR 2

IF2123 ALJABAR LINIER DAN GEOMETRI

Image Retrieval dan Music Information Retrieval Menggunakan PCA dan Vektor



Disusun oleh:

Kelompok 12

| | |
|---------------------|----------|
| Dave Danniell Yanni | 13523003 |
|---------------------|----------|

| | |
|---------------------|----------|
| Muhammad Dicky Isra | 13523071 |
|---------------------|----------|

| | |
|-----------------------|----------|
| Muhammad Fathur Rizky | 13523105 |
|-----------------------|----------|

**PROGRAM STUDI TEKNIK INFORMATIKA
SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA
INSTITUT TEKNOLOGI BANDUNG
JL. GANESA 10, BANDUNG 40132**

2024

DAFTAR ISI

| | |
|----------------------------------|-----------|
| DAFTAR ISI | 2 |
| BAB 1 | 3 |
| DESKRIPSI MASALAH | 3 |
| BAB 2 | 4 |
| TEORI SINGKAT | 4 |
| 2.1 Image Retrieval | 4 |
| 2.2 Music Information Retrieval | 4 |
| 2.3 Principal Component Analysis | 4 |
| 2.4 Singular Value Decomposition | 5 |
| BAB 3 | 15 |
| IMPLEMENTASI PROGRAM | 15 |
| 3.1 Arsitektur Program Front-end | 15 |
| 3.2 Arsitektur Program Back-end | 15 |
| BAB 4 | 17 |
| EKSPERIMEN | 17 |
| 4.1 Eksperimen Prediksi | 17 |
| 4.2 Eksperimen Upload Dataset | 18 |
| 4.3 Eksperimen Menghapus Dataset | 19 |
| 4.4 Eksperimen Memutar Lagu | 20 |
| BAB 5 | 22 |
| KESIMPULAN | 22 |
| 5.1 Kesimpulan | 22 |
| 5.2 Saran | 22 |
| 5.3 Refleksi | 22 |
| LAMPIRAN | 24 |
| Referensi | 24 |
| Tautan Repository | 24 |

BAB 1

DESKRIPSI MASALAH

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.

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.

BAB 2

TEORI SINGKAT

2.1 Image Retrieval

Image Retrieval adalah proses pencarian dan pengambilan citra dari kumpulan data besar berdasarkan ciri-ciri tertentu dari citra yang dicari. Ada dua pendekatan utama dalam image retrieval: text-based dan content-based image retrieval (CBIR). Pendekatan text-based menggunakan metadata seperti kata kunci dan deskripsi untuk menemukan citra, sedangkan CBIR memanfaatkan fitur visual seperti warna, tekstur, dan bentuk untuk melakukan pencocokan secara langsung. Dalam CBIR, citra diwakili oleh penanda unik yang terdiri dari fitur-fitur yang diambil dari nilai pixel-nya, memungkinkan sistem untuk membandingkan dan mencocokkan fitur-fitur ini untuk menemukan kesamaan visual

2.2 Music Information Retrieval

Music Information Retrieval (MIR) adalah bidang penelitian yang berfokus pada pengambilan informasi dari data musik. Ini mencakup teknik untuk mengidentifikasi, mengklasifikasikan, dan menemukan musik berdasarkan konten audio atau metadata. MIR menggabungkan berbagai disiplin ilmu, termasuk pemrosesan sinyal audio, pembelajaran mesin, dan teori informasi. Tujuan utamanya adalah untuk memungkinkan pengguna menemukan musik yang relevan dengan cepat dan efisien melalui analisis fitur-fitur musik seperti melodi, ritme, dan harmoni.

2.3 Principal Component Analysis

Principal Component Analysis (PCA) adalah teknik statistik yang digunakan untuk mereduksi dimensi data sambil mempertahankan sebanyak mungkin informasi asli. PCA mengubah data berdimensi tinggi menjadi beberapa dimensi yang lebih kecil, disebut principal components, tanpa kehilangan esensi atau pola utama dalam data tersebut. Proses ini dilakukan dengan menghitung eigenvector dan nilai eigen dari matriks kovarians data, sehingga memungkinkan identifikasi pola dan tren dalam dataset yang kompleks. PCA sangat berguna dalam berbagai aplikasi, termasuk pengenalan pola dan pemrosesan sinyal karena membantu menyederhanakan analisis data dengan mengurangi jumlah variabel yang perlu dipertimbangkan.

2.4 Singular Value Decomposition

Singular Value Decomposition (SVD) adalah metode matematis yang digunakan untuk menganalisis matriks dengan memecahnya menjadi tiga matriks: satu matriks diagonal yang berisi nilai singular dan dua matriks ortogonal. SVD sering digunakan dalam pemrosesan citra dan pengenalan pola karena memungkinkan reduksi dimensi serta identifikasi struktur dalam data. Dengan SVD, informasi penting dapat diekstraksi dari dataset besar sambil menghilangkan noise atau variabel yang tidak signifikan. Metode ini juga berperan penting dalam rekomendasi sistem dan analisis data multivariat.

BAB 3

IMPLEMENTASI PROGRAM

3.1 Arsitektur Program Front-end

Struktur aplikasi front-end ini dirancang secara komprehensif dengan menggunakan berbagai teknologi modern untuk memastikan pengembangan yang efisien dan skalabilitas yang baik. Next.js dipilih sebagai framework utama karena kemampuannya dalam membangun aplikasi web yang cepat dan dinamis, sementara Tailwind CSS digunakan untuk styling dengan pendekatan utility-first, memungkinkan desain yang responsif dan fleksibel. Validasi form dilakukan dengan Zod, yang memberikan cara yang aman dan deklaratif untuk memverifikasi data input. LucideReact menyediakan ikon-ikon yang modern dan mudah diintegrasikan, meningkatkan pengalaman pengguna dengan visual yang konsisten dan intuitif. Zustand digunakan untuk mengelola state global aplikasi secara efisien, dengan pendekatan yang mendukung pengembangan secara agile dan memastikan integrasi yang lancar antar komponen. ShadCN digunakan untuk membangun komponen-komponen yang dapat digunakan kembali, mempercepat pengembangan dan menjaga konsistensi antarmuka pengguna di seluruh aplikasi. Gabungan teknologi ini memungkinkan pembuatan aplikasi yang terstruktur, mudah dipelihara, dan responsif terhadap kebutuhan pengguna.

3.2 Arsitektur Program Back-end

Struktur umum aplikasi backend ini didesain dengan menggunakan Flask sebagai framework utama. Flask memungkinkan pembuatan aplikasi web dengan cepat dan mudah, serta mendukung pengembangan RESTful API. Selain itu, aplikasi ini juga menggunakan CORS (Cross-Origin Resource Sharing) untuk mengizinkan permintaan dari domain lain, hal ini sangat penting bagi aplikasi web modern yang sering berinteraksi dengan frontend yang terpisah.

Dalam implementasinya, aplikasi ini memiliki dua model utama: ImageModel dan AudioModel. Kedua model ini bertugas untuk memproses dan menganalisis data gambar dan audio secara spesifik. Model-model ini dipanggil setiap kali ada permintaan yang berkaitan dengan proses analisis media.

Untuk mengatur konfigurasi dasar, aplikasi ini membaca informasi dari file settings.json. Konfigurasi ini mencakup detail penting seperti lokasi folder penyimpanan file audio dan gambar. Hal ini memudahkan pengelolaan data dan integrasi dengan model yang sudah ada.

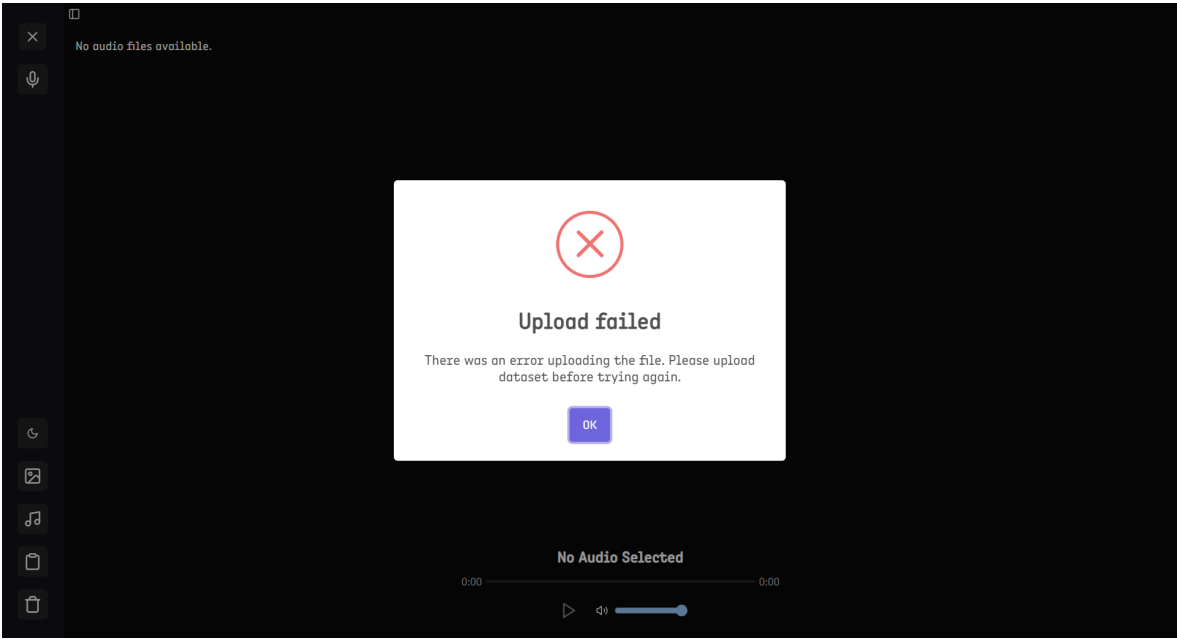
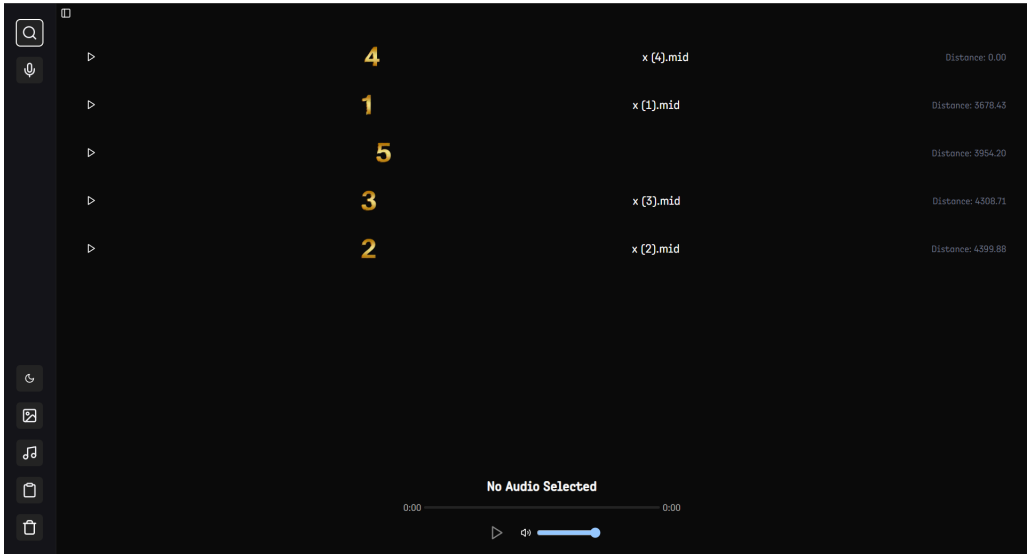
Tabel 1. Manajemen Endpoint API

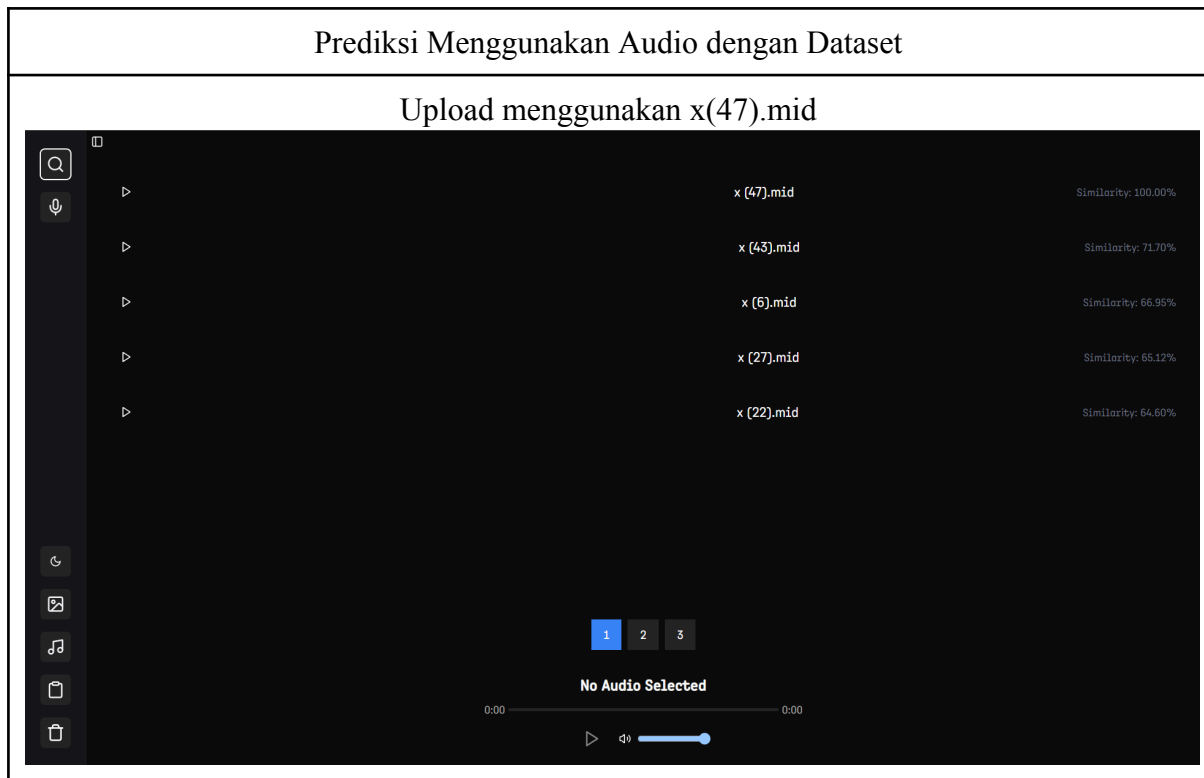
| Kategori | Endpoint | Metode HTTP | Deskripsi |
|----------------------------|---------------|-------------|---|
| Manajemen Mapper | get/mapper | GET | Mengambil mapper yang aktif |
| | upload/mapper | POST | Mengunggah file .json atau .txt untuk memperbarui mapper. |
| Manajemen Audio dan Gambar | upload/audio | POST | Mengunggah file audio, baik file individual maupun file terkompresi. |
| | upload/image | POST | Mengunggah file gambar, baik file individual maupun file terkompresi. |
| Prediksi Audio dan Gambar | predict/audio | POST | Memprediksi audio yang mirip dari database berdasarkan query file. |
| | predict/image | POST | Memprediksi gambar yang mirip dari database berdasarkan query file. |
| Reset API | reset | GET | Menghapus seluruh data dari direktori database dan mengatur ulang mapper. |
| Health Check | health | GET | Memastikan server sudah berjalan dengan baik |

BAB 4

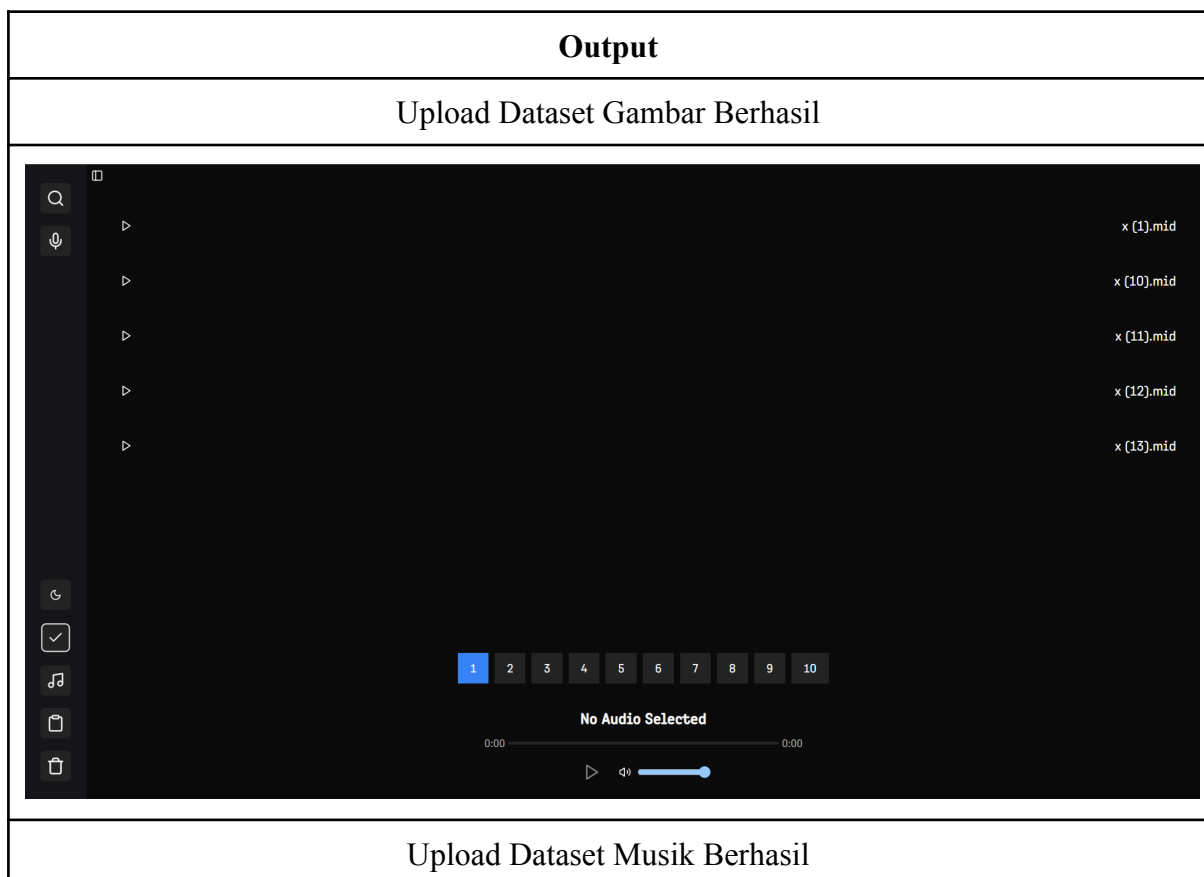
EKSPERIMEN

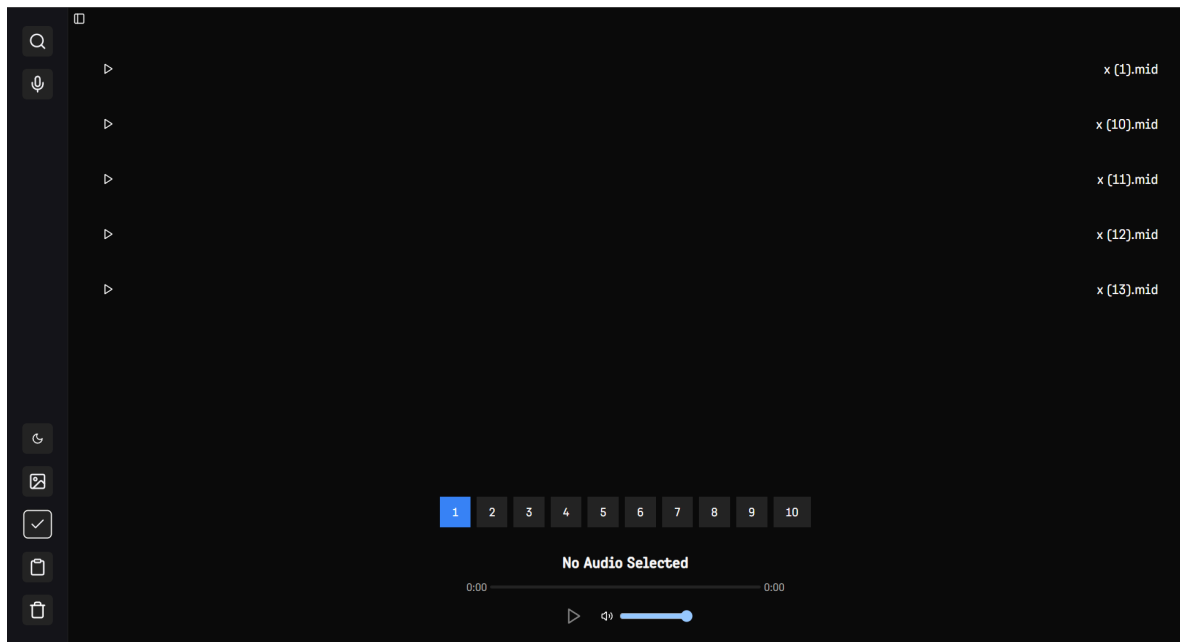
4.1 Eksperimen Prediksi

| Output |
|--|
| <p>Prediksi Tanpa Dataset</p>  |
| <p>Prediksi Menggunakan Gambar dengan Dataset</p> <p>Prediksi menggunakan gambar 4</p>  |

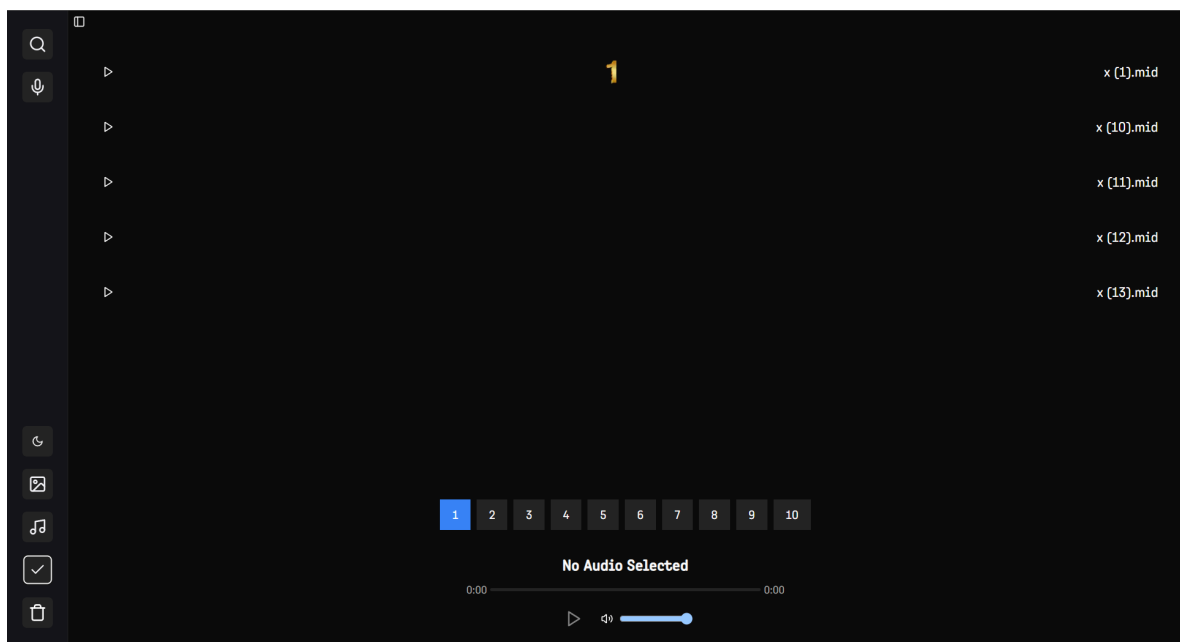


4.2 Eksperimen Upload Dataset



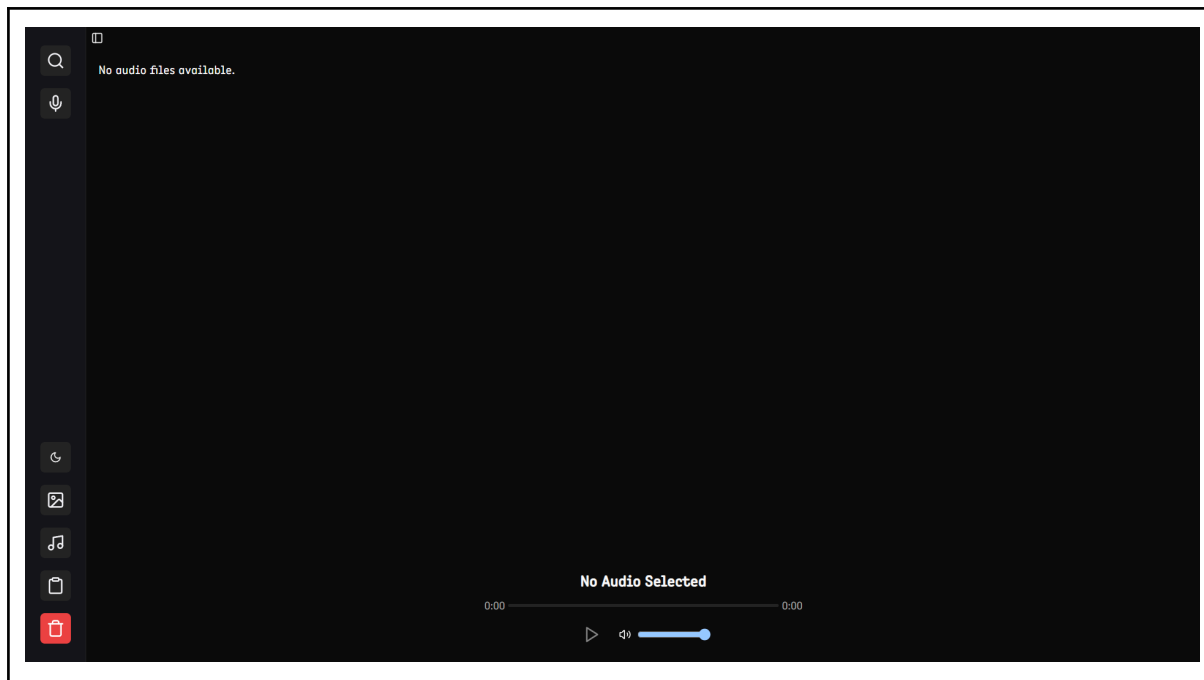


Upload Dataset Mapper Berhasil

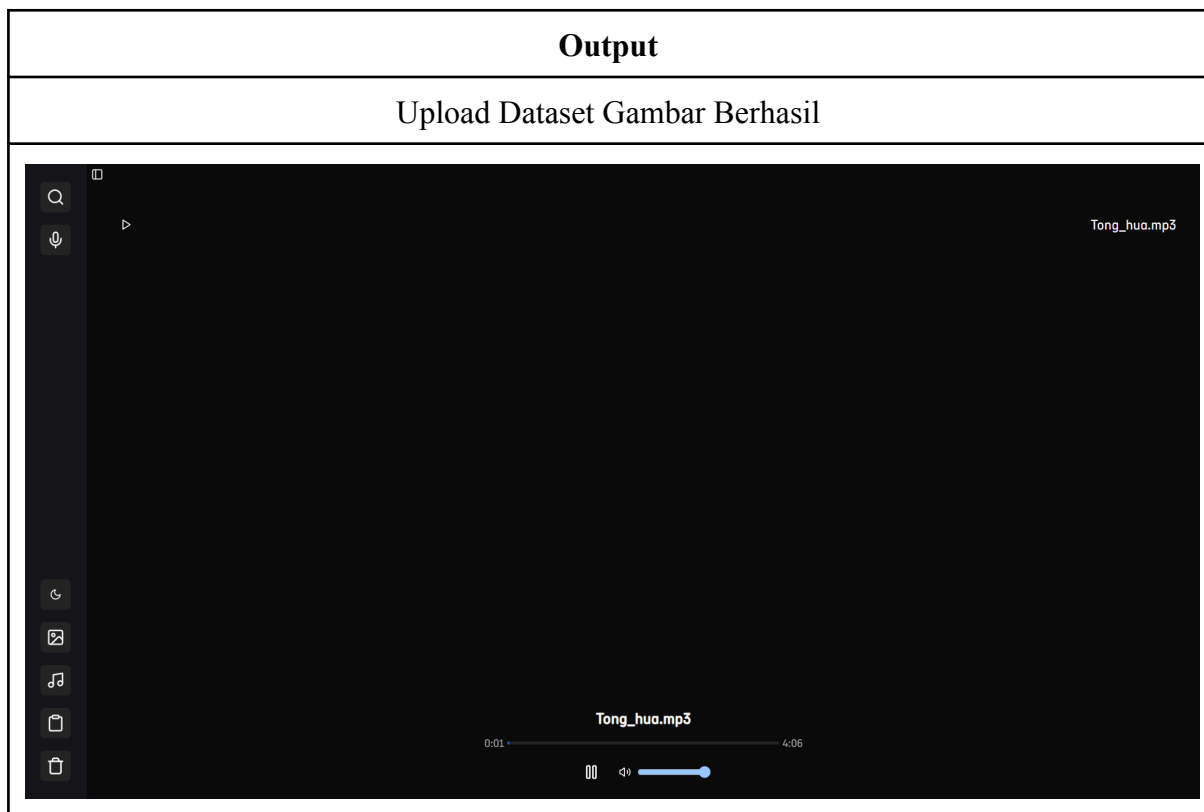


4.3 Eksperimen Menghapus Dataset

| Output |
|-------------------|
| Menghapus Dataset |



4.4 Eksperimen Memutar Lagu



BAB 5

KESIMPULAN

5.1 Kesimpulan

Pada Tugas Besar 2 ini, kami mengembangkan sebuah website untuk Information Retrieval yang mengintegrasikan konsep matematika dalam teknologi pencarian multimedia. Fokus utama kami adalah mengimplementasikan dua metode inovatif: Music Information Retrieval (Query by Humming) dan Album Finder dengan menggunakan Principal Component Analysis (PCA). Dalam metode Music Information Retrieval, kami mengolah file MIDI untuk mencocokkan melodi melalui ekstraksi fitur kompleks seperti distribusi nada, dengan menggunakan cosine similarity untuk menghitung kemiripan. Sementara itu, dalam Album Finder, kami menerapkan teknik PCA untuk mentransformasi gambar menjadi grayscale, melakukan standarisasi data, dan menghitung kedekatan gambar menggunakan jarak Euclidean. Website yang kami kembangkan mencakup fitur-fitur kunci seperti unggah dataset, pagination, perhitungan persentase kemiripan, dan pencatatan waktu eksekusi. Kami diperbolehkan menggunakan query berbasis audio atau gambar, dengan kebebasan untuk memilih teknologi frontend dan backend yang paling sesuai dengan kemampuan kami. Selain tugas utama, kami juga dapat mengembangkan fitur bonus untuk meningkatkan nilai, seperti video name mapping, benchmark performa, dan fitur kreativitas tambahan.

5.2 Saran

Pada tugas besar 2 ini, dapat dilakukan pengembangan seperti mikrofon real-time yang dapat berjalan dengan baik.

5.3 Refleksi

Pada Tugas Besar 2 ini, kami mengasah kemampuan mendekomposisi masalah dan pemrograman dengan menggunakan Python Flask serta meningkatkan pemahaman mengenai penerapan prinsip eigenvector, khususnya dengan menggunakan teknik PCA (Principal Component Analysis) dan SVD (Singular Value Decomposition). Selain itu, proyek ini juga memperkenalkan penggunaan framework modern, yakni Next.js dengan ShadCN pada

frontend untuk membangun aplikasi yang mengintegrasikan teori aljabar linier dalam bentuk yang interaktif.

Melalui tugas besar ini, kami tidak hanya belajar mengimplementasikan teori PCA dan SVD dalam aplikasi praktis, tetapi juga mengembangkan kemampuan dalam kolaborasi tim dan komunikasi antar anggota, sambil memperdalam pemahaman mengenai aplikasi aljabar linier dalam teknologi pencarian multimedia, seperti dalam konteks Music Information Retrieval dan Album Finder.

..

LAMPIRAN

Referensi

Abdi, H., & Williams, L. J. (2010). Principal Component Analysis. In Wiley Interdisciplinary Reviews: Computational Statistics, 2(4), 433-459.

Datta, R., Joshi, D., Li, J., & Wang, J. Z. (2008). Image Retrieval: Ideas, Influences, and Trends of the New Age. ACM Computing Surveys, 40(2), 1-60.

Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A., & Jain, R. (2000). Content-Based Image Retrieval at the End of the Early Years. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(12), 1349-1380.

Jolliffe, I. T. (2002). Principal Component Analysis. Springer Series in Statistics.

Golub, G. H., & Van Loan, C. F. (1996). Matrix Computations (3rd ed.). Johns Hopkins University Press.

De Silva, V., & Tenenbaum, J. B. (2004). Sparse Multidimensional Scaling Using Landmark Points. In Proceedings of the National Academy of Sciences, 101(30), 10308-10313.

Tautan *Repository*

<https://github.com/fathurwithyou/Algeo02-23003>