

LAPORAN
IF2123 Aljabar Linier dan Geometri
Tugas Besar 2
Semester I 2025/2026



Disusun oleh:

Kelompok 10 LemonNipis
Niko Samuel Simanjuntak — 13524029
Irvin Tandiarrang Sumual — 13524030
Kalyca Nathania B. Manullang — 13524071

Institut Teknologi Bandung

2025

1 PENDAHULUAN

1.1 Kisah

Di Desa Konohagakure, hiduplah dua orang mahasiswa Teknik Informatika bernama Ayu dan Asep Kasep. Akhir-akhir ini, Ayu merasa kewalahan memahami materi dekomposisi matriks yang diajarkan dalam mata kuliah IF2123 Aljabar Linier dan Geometri karena ketimbang memperhatikan penjelasan dosen di depan, ia malah asyik *doomscrolling* di media sosial. Supaya bisa bertanggung jawab kepada kedua orang tuanya yang telah membiayai kuliahnya, Ayu pun bertekad untuk memperbaiki situasi dengan belajar ke perpustakaan. Sesampainya di sana, ia takjub akan banyaknya koleksi buku yang tersedia. Sejak saat itu, Ayu pun gemar mengunjungi perpustakaan untuk mengeksplorasi berbagai macam buku yang menarik minatnya.

Di sisi lain, Asep adalah seorang mahasiswa yang bekerja paruh waktu sebagai pus-takawan di perpustakaan kampus. Sejak Ayu mulai rajin berkunjung ke perpustakaan, Asep sering membantunya mencari buku-buku yang diinginkannya. Awalnya, Ayu mengira Asep adalah seorang *performative male* yang hanya ingin pamer di hadapannya. Apalagi, Asep sering sekali minum matcha. Namun, seiring berjalannya waktu, Ayu menyadari bahwa Asep adalah sosok yang cerdas dan tulus dalam membantu sesama. Asep juga sadar bahwa Ayu adalah pribadi yang tekun dan pantang menyerah. Akhirnya, mereka pun jatuh cinta.



Figure 1: Anggaplah ini Ayu dan Asep.

Sumber: Jaap Buitendijk; Copyright 2007 Warner Bros. Ent.; Harry Potter Publishing Rights J.K.R.

Asep menyadari bahwa Ayu sering kali kesulitan menemukan buku yang diinginkannya atau melanjutkan bacaan dari buku yang pernah dipinjamnya sebelumnya. Hal ini dikarenakan sistem pencarian buku di perpustakaan yang masih konvensional dan kurang efisien. Melihat hal tersebut, Asep terinspirasi untuk mengembangkan sebuah sistem temu balik informasi berbasis *singular value decomposition* (SVD) dan vektor eigen yang dapat membantu Ayu dan pengunjung perpustakaan lainnya dalam menemukan buku dengan lebih mudah dan cepat.

Untuk membuktikan bahwa mahasiswa Teknik Informatika memanglah solid, mari kita bantu Asep dalam merancang dan mengimplementasikan sistem temu balik informasi tersebut dengan memanfaatkan konsep-konsep aljabar linier yang telah kita pelajari di mata kuliah IF2123 Aljabar Linier dan Geometri.

1.2 Tujuan

Pada Tugas Besar 2 ini, kita akan mengembangkan sebuah aplikasi web untuk sistem perpustakaan digital. Anda akan diberikan sebuah dataset berisi koleksi buku dengan rincian judul, konten (berupa file txt), dan sampul (berupa file jpg). Aplikasi ini harus mampu menampilkan daftar buku dalam dataset dengan paginasi, menampilkan konten buku untuk dibaca, melakukan pencarian buku berdasarkan input sampul gambar dari pengguna atau kata kunci teks dari pengguna, serta memberikan rekomendasi buku yang relevan berdasarkan buku yang sedang dibaca.

2 DASAR TEORI

2.1 Nilai Eigen dan Vektor Eigen

Diberikan matriks persegi A berukuran $n \times n$, vektor tak-nol \mathbf{x} disebut vektor eigen jika $A\mathbf{x} = \lambda\mathbf{x}$, dengan λ sebagai nilai eigen. Vektor eigen hanya mengalami perubahan skala saat ditransformasi, merepresentasikan arah karakteristik transformasi linear. Nilai eigen dicari dari persamaan karakteristik $\det(A - \lambda I) = 0$, lalu vektor eigen diperoleh dari $(A - \lambda I)\mathbf{x} = \mathbf{0}$. Secara geometris, eigenvector menunjukkan arah dominan transformasi, sedangkan eigenvalue menentukan besar kontribusi (varians atau energi).

2.2 Singular Value Decomposition (SVD)

Setiap matriks A berukuran $m \times n$ dapat didekomposisi menjadi $A = U\Sigma V^T$, dengan U dan V matriks ortogonal, serta Σ matriks diagonal berisi singular values $\sigma_1 \geq \sigma_2 \geq \dots$. Vektor-vektor pada V adalah eigenvector dari $A^T A$ dengan σ_i^2 sebagai eigenvalue-nya. SVD memberikan landasan kuat untuk ekstraksi fitur dari teks dan gambar.

2.3 Hubungan Eigenvector dengan Representasi Fitur Utama

Matriks kovarian $C = \frac{1}{n}XX^T$ merepresentasikan penyebaran data. Eigenvalue menunjukkan besar varians pada arah eigenvector. Arah dengan eigenvalue terbesar merupakan fitur paling informatif, sedangkan arah dengan eigenvalue kecil mengandung noise. Eigenvector dari matriks kovarian atau dari $A^T A$ melalui SVD menangkap pola dominan seperti komponen bentuk pada gambar atau hubungan kata pada teks.

2.4 Mengapa SVD Lebih Cocok daripada Eigendecomposition

SVD lebih cocok digunakan daripada eigendecomposition untuk tugas besar ini karena tiga alasan utama. Pertama, SVD dapat bekerja pada matriks tidak persegi dan berukuran besar, sedangkan eigendecomposition hanya berlaku untuk matriks persegi. Dataset

teks dan gambar dalam tugas ini berbentuk matriks dokumen-kata atau gambar flatten yang tidak persegi, sehingga fleksibilitas SVD menjadi keunggulan. Kedua, SVD lebih stabil secara numerik dan robust terhadap noise, sementara menghitung eigen dari matriks $A^T A$ dapat memperbesar kesalahan numerik karena pemangkatan nilai besar dan dominansi noise. Ketiga, SVD menghindari komputasi matriks kovarian $A^T A$ yang mahal dan tidak stabil untuk dataset besar. Dengan SVD, fitur utama dapat diperoleh langsung dari dekomposisi matriks A tanpa perlu menghitung matriks kovarian terlebih dahulu, sehingga lebih efisien dan andal.

2.5 Mengapa Eigenvector dari Matriks Covariance Merepresentasikan Fitur Utama

Matriks kovarian mengukur varians data pada setiap arah. Eigenvector dengan eigenvalue terbesar menunjukkan arah dengan penyebaran data terbesar sehingga menjadi fitur paling informatif. Oleh karena itu, PCA menggunakan eigenvector kovarian sebagai principal components untuk menangkap variasi terpenting dalam data dengan dimensi lebih rendah.

2.6 Kapan PCA Berbasis SVD Kurang Optimal untuk Kueri Gambar

PCA berbasis SVD menjadi kurang optimal untuk kueri gambar dalam beberapa kondisi. Pertama, PCA sensitif terhadap perbedaan pencahayaan ekstrem karena perubahan brightness memengaruhi varians global, sehingga gambar dengan objek sama tetapi pencahayaan berbeda dapat direpresentasikan secara jauh berbeda. Kedua, PCA memiliki keterbatasan terhadap transformasi geometris seperti rotasi, translasi, atau scaling karena bekerja pada representasi piksel linear; transformasi tersebut menggeser posisi piksel dan membuat eigenvector tidak konsisten. Ketiga, metode ini cenderung mengabaikan detail kecil (seperti kontur tipis atau tekstur halus) yang memiliki varians rendah karena lebih berfokus pada pola global dengan varians besar. Keempat, pada gambar yang sangat noisy atau bertekstur kompleks, noise dengan varians tinggi dapat mendominasi principal components dan mengalihkan fokus PCA. Untuk kasus-kasus tersebut, metode berbasis fitur lokal seperti SIFT, ORB, atau CNN sebenarnya lebih cocok, namun tidak digunakan dalam tugas besar ini karena keterbatasan ruang lingkup.

3 DESAIN DAN IMPLEMENTASI

3.1 Arsitektur Aplikasi

Aplikasi dirancang dengan arsitektur full-stack yang memisahkan frontend (Next.js) dan backend (FastAPI). Terdiri dari empat komponen: frontend untuk antarmuka pengguna, backend untuk menangani permintaan, model dan dataset berisi data buku dan model praproses, serta modul pemrosesan data (LSA untuk teks, substring matching untuk judul, PCA untuk gambar). Alur sistem: **Dataset → Preprocessing → Model → Query → Hasil**.

3.2 Teknologi yang Digunakan

| Teknologi | Peran |
|-----------------|------------------------------------|
| Backend | |
| Python 3.x | Bahasa utama backend |
| FastAPI | Framework API |
| NumPy | Operasi matriks (TF-IDF, SVD, PCA) |
| Regex (re) | Tokenisasi teks |
| NLTK | Stemming dan stopword removal |
| Pickle | Penyimpanan model LSA |
| PIL / OpenCV | Pengolahan gambar |
| Frontend | |
| Next.js 14 | Framework React |
| HeroUI v2 | Komponen UI |
| Tailwind CSS | Styling |
| TypeScript | Tipe statis |
| Framer Motion | Animasi |

Table 1: Teknologi yang Digunakan

3.3 Struktur Program dan Modul-Modul Penting

```

root/
  data/
  docs/
  src/
    backend/
      document/
        document_processing.py
      image/
        image_processing.py
        tempCodeRunnerFile.py
        __init__.py
      model/
        build_model.py
      .gitkeep
      main.py
    frontend/
      .vscode/
        settings.json
    app/
      book-collection/
        [id]/
          page.tsx
        layout.tsx
        page.tsx
  
```

```
search-result/
    page.tsx
    error.tsx
    layout.tsx
    page.tsx
    providers.tsx
components/
    book-detail/
config/
    fonts.ts
    site.ts
public/
    favicon.ico
    next.svg
    vercel.svg
styles/
    global.css
types/
    index.ts
.gitignore
.gitkeep
.npmrc
LICENSE
README.md
eslint.config.mjs
next.config.js
package.json
postcss.config.js
tailwind.config.js
tsconfig.json
test/
    .gitkeep
    .gitignore
LICENSE
README.md
```

3.3.1 Modul document_processing.py

Inti pencarian berbasis teks dengan LSA: tokenisasi, normalisasi, stopword removal, TF-IDF, truncated SVD, cosine similarity. Menyediakan `query_lsa()` dan `load_lsa_model()`.

3.3.2 Modul build_model.py

Script offline untuk membangun model LSA: memuat dataset, preprocessing, TF-IDF, SVD, menyimpan ke `model.pkl`.

3.3.3 Modul image_processing.py

Pemrosesan gambar: grayscale, flatten vektor, PCA/SVD, proyeksi query, cosine similarity.

3.3.4 Modul main.py

Pusat integrasi backend: inisialisasi model, API endpoint, pemetaan input, format respons.

3.3.5 Frontend

Dibangun dengan Next.js: halaman utama, daftar buku, detail buku, pencarian. Komponen: Navbar, SearchBar, BookCard. Komunikasi dengan backend via fetch API.

3.4 Cara Kerja dan Alur Program

3.4.1 Inisialisasi

Backend memuat model LSA, embedding gambar PCA, dan metadata buku sekali saat startup.

3.4.2 Input dari Pengguna

Frontend menyediakan tiga mode input: pencarian judul, dokumen, atau gambar.

3.4.3 Penentuan Metode Pencarian

Backend mengarahkan ke fungsi yang sesuai: title_search() untuk judul, query_lsa() untuk dokumen, atau PCA untuk gambar.

3.4.4 Proses Pencarian

- **Dokumen (LSA):** Preprocess → TF-IDF → proyeksi ruang laten → cosine similarity.
- **Judul:** Substring matching.
- **Gambar:** Proyeksi ke PCA eigen-space → cosine similarity dengan embedding gambar lain.

4 EKSPERIMEN

4.1 Kasus Uji Pencarian Berdasarkan Judul

4.1.1 Kasus Uji 1: Pencarian "American"

- **Query:** "American"

- **Hasil:** Sistem berhasil menemukan buku dengan judul yang mengandung kata "American".

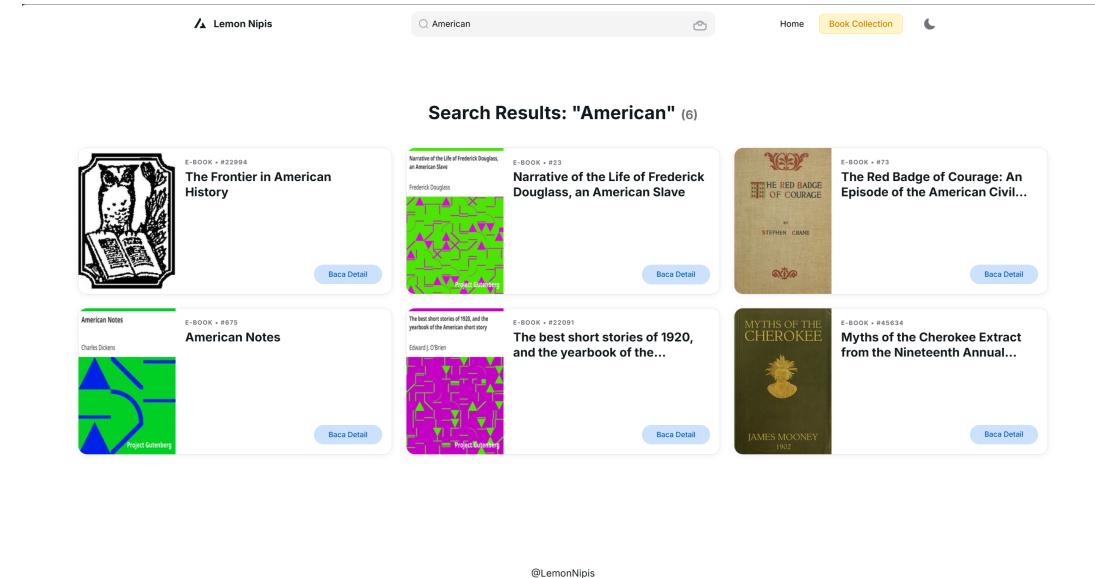


Figure 2: Hasil pencarian untuk judul "American"

4.1.2 Kasus Uji 2: Pencarian "Sherlock"

- **Query:** "Sherlock"
- **Hasil:** Sistem berhasil menemukan buku-buku dengan judul yang mengandung "Sherlock".

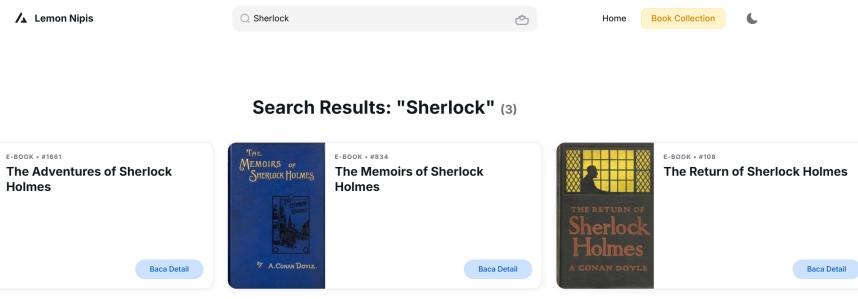


Figure 3: Hasil pencarian untuk judul "Sherlock"

4.1.3 Kasus Uji 3: Pencarian "The World as Will and Idea (Vol. 1 of 3)"

- **Query:** "The World as Will and Idea (Vol. 1 of 3)"
- **Hasil:** Sistem berhasil menemukan buku dengan judul yang sesuai.

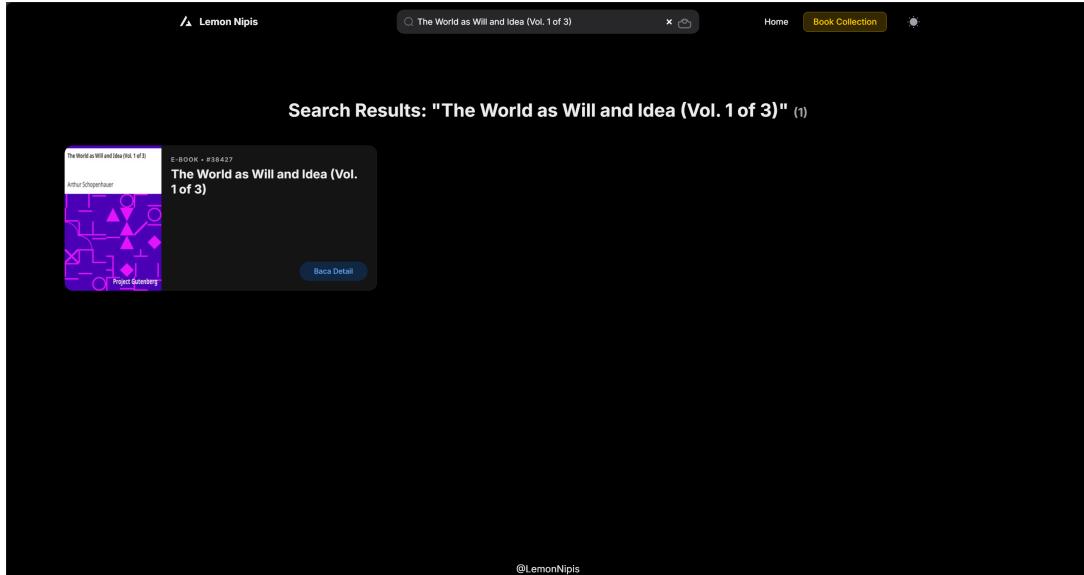


Figure 4: Hasil pencarian untuk judul "The World as Will and Idea (Vol. 1 of 3)"

4.2 Kasus Uji Pencarian Berdasarkan Topik/Keyword

4.2.1 Kasus Uji 1: Pencarian Dokumen "The Adventures of Sherlock Holmes" (ID : 108)

- **Query:** Dokumen teks berisi awal buku *The Adventures of Sherlock Holmes*.
- **Hasil:** Sistem LSA berhasil menemukan buku-buku dengan tingkat kemiripan konten yang sangat tinggi. Buku dengan *Book ID 108* (*The Return of Sherlock Holmes*) muncul sebagai hasil paling relevan dengan skor similarity 0.9846, diikuti oleh beberapa buku lain dalam seri Sherlock Holmes yang juga memiliki skor di atas 0.98.

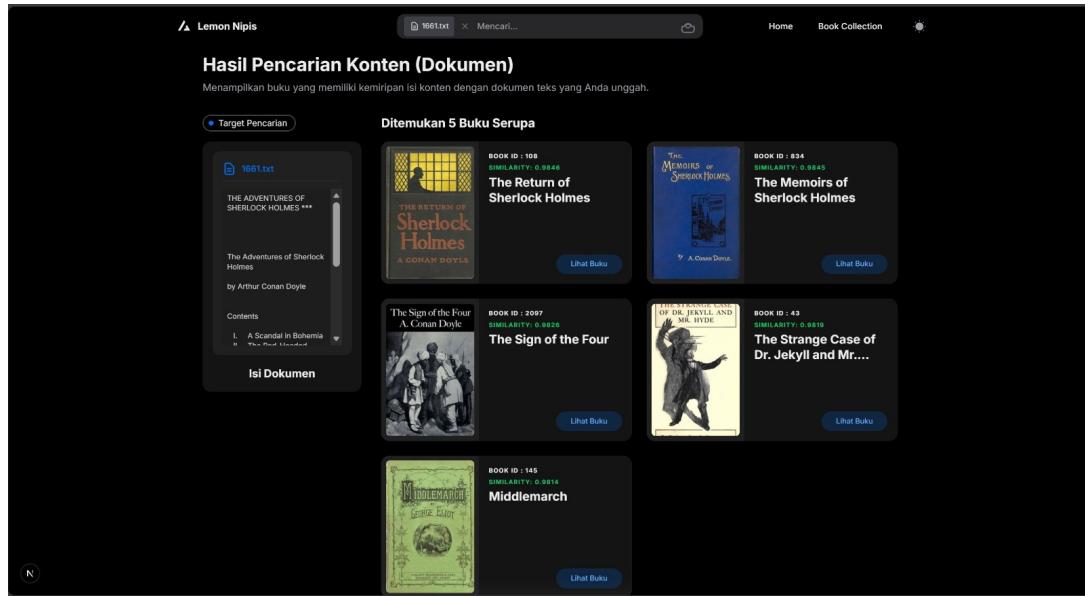


Figure 5: Hasil pencarian untuk dokumen "The Adventures of Sherlock Holmes"

4.2.2 Kasus Uji 2: Pencarian Dokumen "THE DECLARATION OF INDEPENDENCE of the united states of america"

- Query:** Dokumen teks berisi awal buku *THE DECLARATION OF INDEPENDENCE of the united states of america*.
- Hasil:** Sistem LSA berhasil mengenali bahwa dokumen input memiliki kemiripan konten yang sangat kuat dengan buku-buku bertema sejarah dan politik Amerika. Buku dengan Book ID 1 ("The Declaration of Independence...") muncul sebagai hasil paling relevan dengan skor similarity 0.9647, selaras dengan isi dokumen input. Empat buku lainnya juga memiliki skor tinggi di rentang 0.92–0.94, menunjukkan bahwa model mampu menangkap keterkaitan semantik antar-teks secara konsisten.

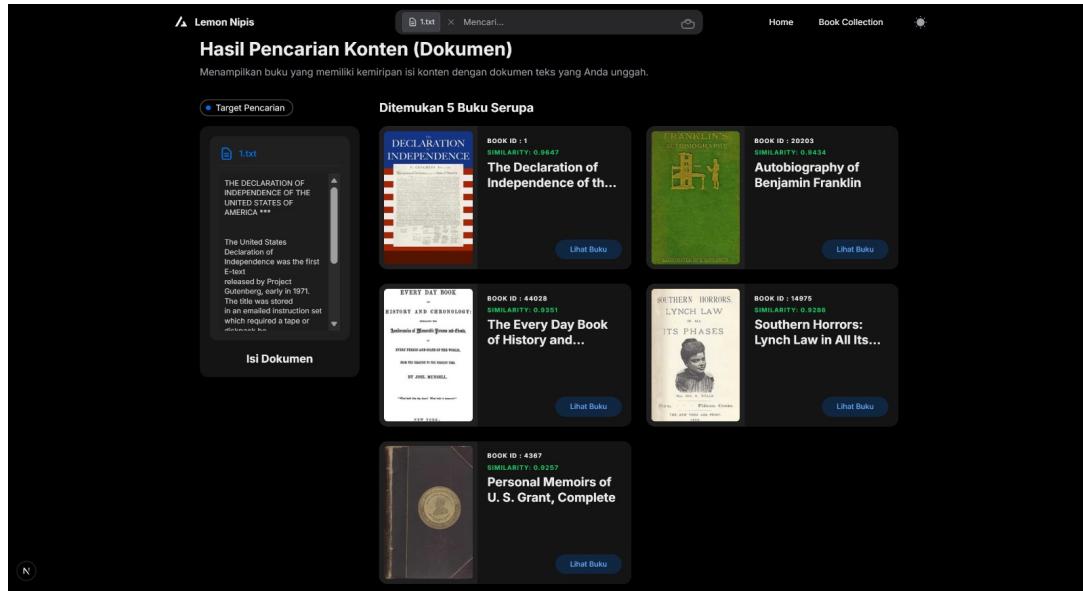


Figure 6: Hasil pencarian untuk dokumen "THE DECLARATION OF INDEPENDENCE of the united states of america"

4.3 Kasus Uji Pencarian Berdasarkan Gambar Sampul Buku dari Asisten

4.3.1 Kasus Uji 1: Gambar Sampul "The Wonderful Wizard of Oz"

- **Hasil:** Sistem berhasil mengenali buku yang benar, ditunjukkan oleh Book ID 55 yang sesuai dengan expected ID dan memiliki skor kemiripan terendah (1641.4558) yang menandakan kemiripan visual tertinggi.

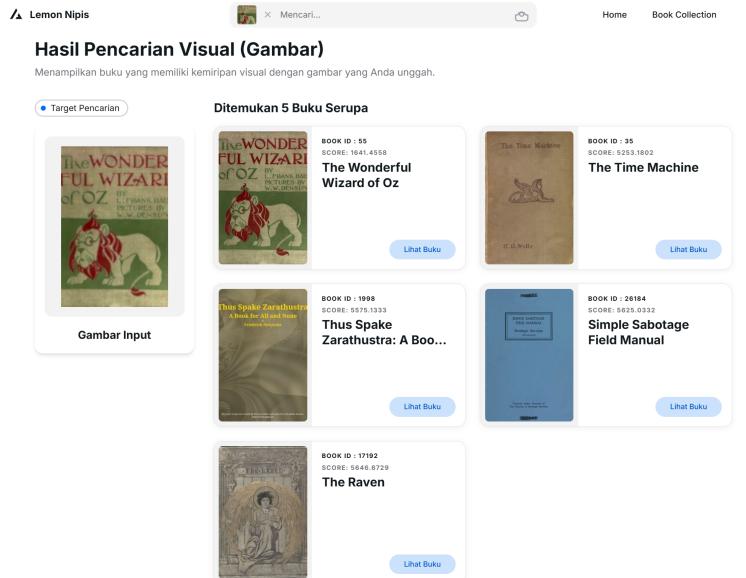


Figure 7: Hasil pencarian untuk gambar sampul "The Wonderful Wizard of Oz"

4.3.2 Kasus Uji 2: Gambar Sampul "A Treatise of Human Nature"

- **Hasil:** Sistem berhasil mengidentifikasi buku yang sesuai, ditunjukkan oleh Book ID 4705 yang sesuai dengan expected ID dan memiliki skor kemiripan terendah (3051.4399) yang menandakan kemiripan visual tertinggi.

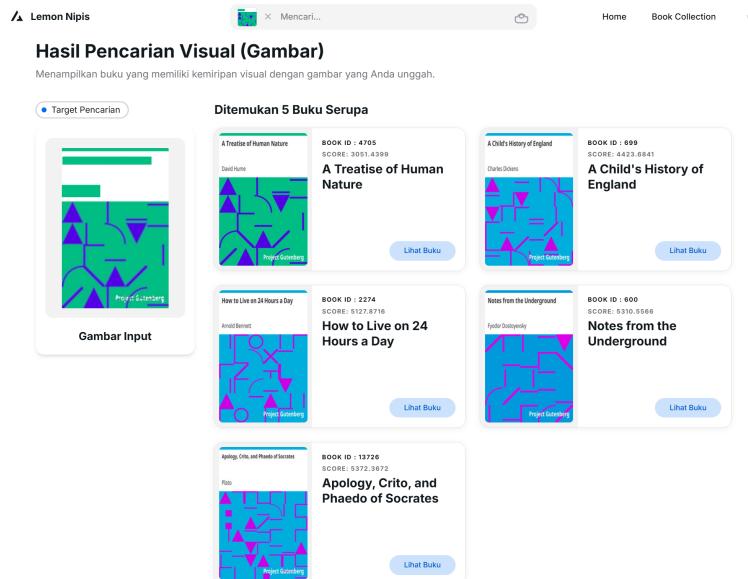


Figure 8: Hasil pencarian untuk gambar sampul "A Treatise of Human Nature"

4.3.3 Kasus Uji 3: Gambar Sampul "The Adventures of Sherlock Holmes"

- **Hasil:** Sistem berhasil mengidentifikasi buku yang sesuai, ditunjukkan oleh Book ID 1661 yang sesuai dengan expected ID dan memiliki skor kemiripan terendah (2283.1819) yang menandakan kemiripan visual tertinggi.

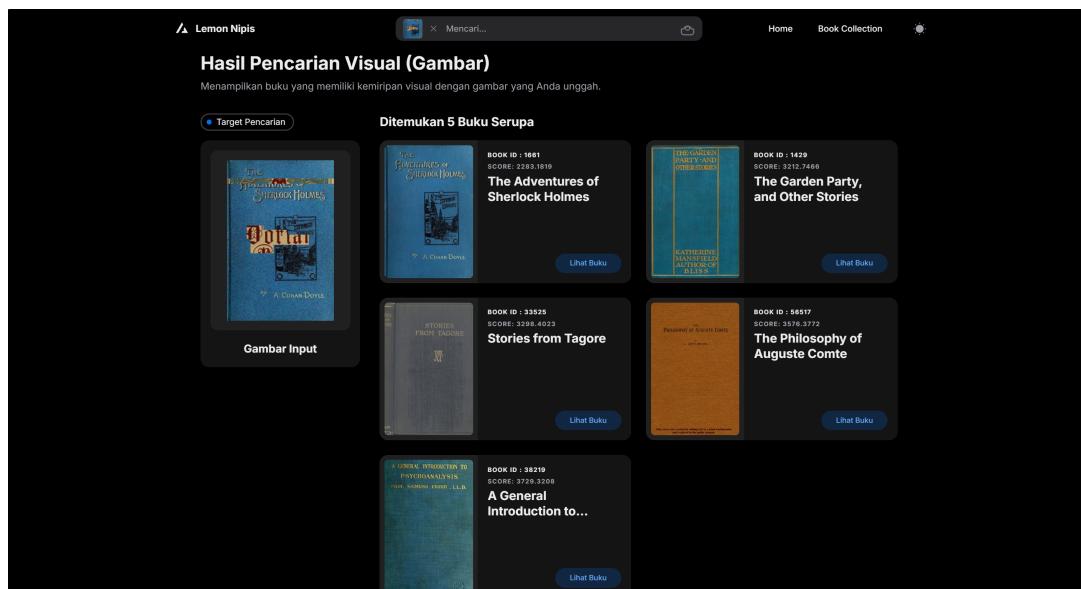


Figure 9: Hasil pencarian untuk gambar sampul "The Adventures of Sherlock Holmes"

4.4 Kasus Uji Pencarian Berdasarkan Gambar Sampul Buku yang Ditentukan Sendiri

4.4.1 Kasus Uji 1: Gambar Sampul ”The Merchant of VENICE”

- **Hasil:** Sistem berhasil mengidentifikasi buku yang benar, ditunjukkan oleh Book ID 1515 yang sesuai dengan expected ID dan memiliki skor kemiripan 0.0228, jauh lebih tinggi dibanding buku lain (tidak ditampilkan karena skornya jauh lebih rendah).

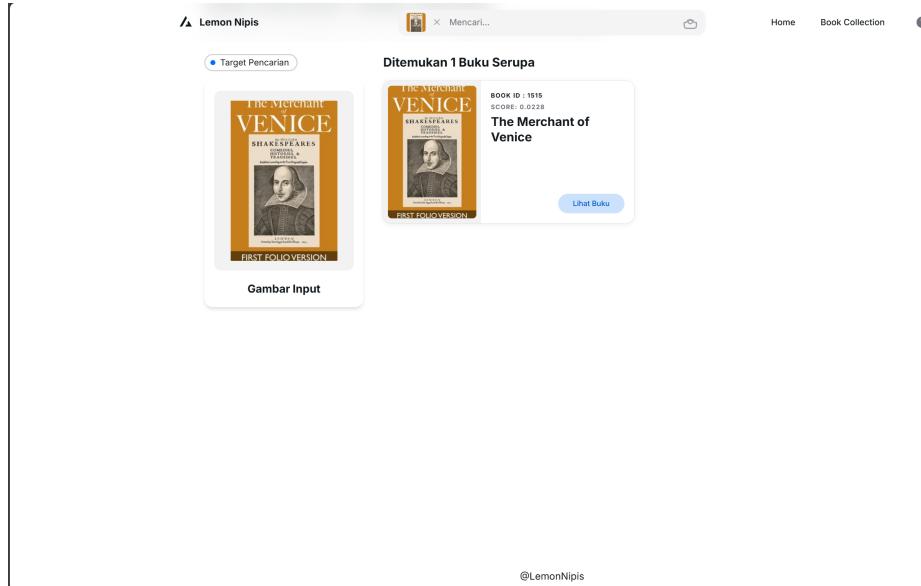


Figure 10: Hasil pencarian untuk gambar sampul ”The Merchant of VENICE”

4.4.2 Kasus Uji 2: Gambar Sampul ”Paradise Lost”

- **Hasil:** Sistem berhasil mengidentifikasi Paradise Lost dengan Book ID 20, sesuai dengan expected ID. Skor kemiripannya (0.0183) merupakan nilai terendah di antara hasil lain, nilai yang lebih kecil menunjukkan kemiripan yang lebih tinggi.

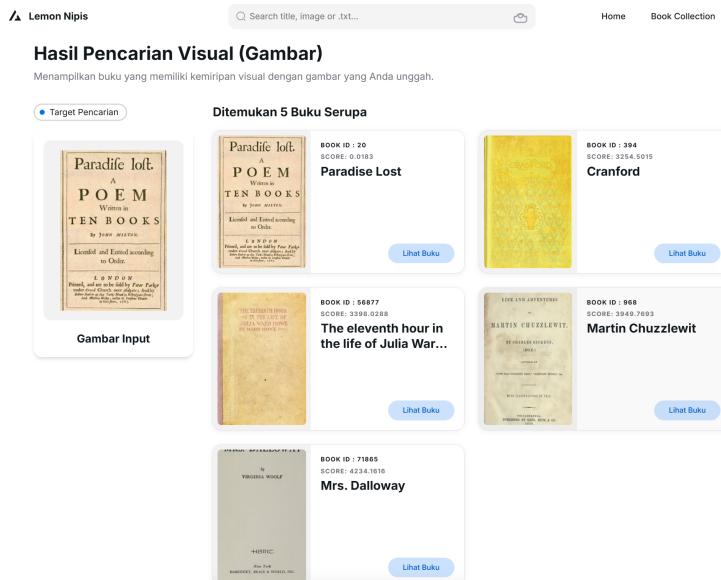


Figure 11: Hasil pencarian untuk gambar sampul "Paradise Lost"

4.5 Tangkapan Layar Rekomendasi Buku

4.5.1 Tangkapan Layar 1: Rekomendasi Setelah Membaca "The Return of Sherlock Holmes (ID:108)"

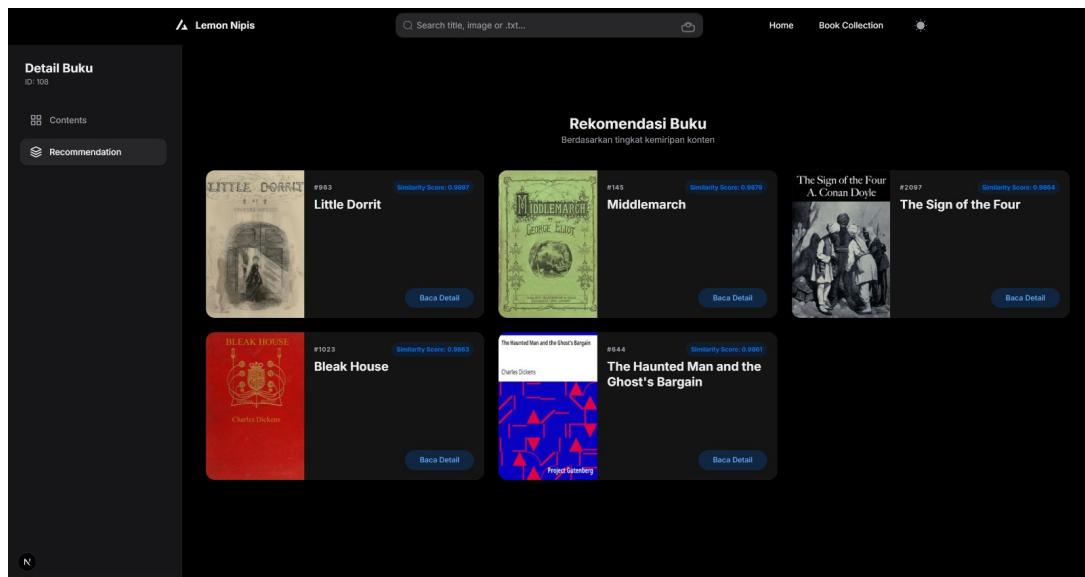


Figure 12: Rekomendasi buku setelah membaca "The Return of Sherlock Holmes (ID:108)"

Hasil Analisis: Sistem berhasil memberikan rekomendasi 5 buku lain yang relevan berdasarkan konten, genre, atau penulis. Buku yang direkomendasikan sebagian besar merupakan karya sastra klasik Inggris dari periode Victoria, termasuk karya Charles Dickens dan penulis yang sama (Arthur Conan Doyle).

4.5.2 Tangkapan Layar 2: Rekomendasi Setelah Membaca "THE DECLARATION OF INDEPENDENCE of the united states of america (ID : 1)"

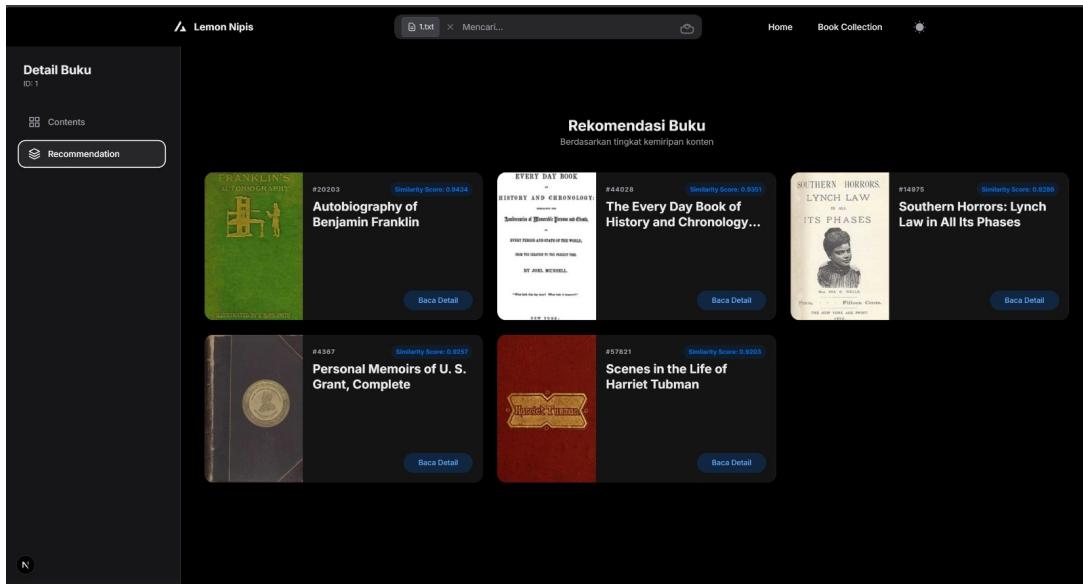


Figure 13: Rekomendasi buku setelah membaca "THE DECLARATION OF INDEPENDENCE of the united states of america (ID : 1)"

Hasil Analisis: Sistem berhasil memberikan rekomendasi buku-buku terkait. Rekomendasi yang diberikan mencakup dokumen-dokumen penting lainnya dalam sejarah Amerika Serikat seperti *The Constitution of the United States* dan *The Federalist Papers*, serta karya-karya pendiri bangsa seperti tulisan Benjamin Franklin.

4.6 Analisis Keseluruhan

Berdasarkan hasil eksperimen, sistem mampu menangani berbagai jenis pencarian dengan baik. Pencarian judul berhasil menemukan buku berdasarkan substring matching. Pencarian dokumen dengan LSA berhasil menemukan buku. Pencarian gambar dengan PCA berhasil mengidentifikasi buku berdasarkan kemiripan visual. Sistem rekomendasi juga berhasil memberikan buku-buku yang relevan.

5 PENUTUP

5.1 Kesimpulan

Tugas Besar 2 Aljabar Linier dan Geometri yang penulis rancang telah berhasil mengembangkan sistem temu balik informasi berbasis SVD dan vektor eigen untuk perpustakaan digital. Aplikasi web yang dibangun mampu melakukan pencarian buku berdasarkan judul (substring matching), topik/keyword dari dokumen (LSA), dan gambar sampul (PCA). Sistem rekomendasi berbasis LSA juga berhasil diimplementasikan untuk merekomendasikan buku serupa berdasarkan konten. Penggunaan SVD terbukti lebih stabil dan

cocok untuk matriks tidak persegi dibandingkan eigendecomposition, sedangkan PCA berbasis SVD efektif untuk ekstraksi fitur gambar.

5.2 Saran kepada Asisten

- Memberikan contoh dataset yang lebih variatif untuk menguji robustnes sistem.
- Menyediakan dokumentasi yang lebih rinci mengenai format input dan output yang diharapkan.
- Memperjelas batasan-batasan teknis yang diizinkan dalam implementasi.

5.3 Refleksi Pribadi

5.3.1 Niko Samuel Simanjuntak (13524029)

Terima kasih asisten atas tugas dan extend deadlinenya.

5.3.2 Irvin Tandiarrang Sumual (13524030)

Terima kasih extend deadlinenya dan tugasnya.

5.3.3 Kalyca Nathania B. Manullang (13524071)

Terima kasih kepada Tuhan Yesus, semangat terus bagi yang membaca laporan ini, semangat Nayaka.

6 DAFTAR PUSTAKA

References

- [1] Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K. dan Harshman, R., "Indexing by latent semantic analysis," J. Amer. Soc. Inf. Sci., vol. 41, no. 6, pp. 391–407, Sept. 1990. [Daring]. Tersedia: https://web.archive.org/web/20241122082607/http://wordvec.colorado.edu/papers/Deerwester_1990.pdf
- [2] J. Shlens, "A Tutorial on Principal Component Analysis," Google Research, Mountain View, CA, Tech. Rep. arXiv:1404.1100v1 [cs.LG], Apr. 2014. [Daring]. Tersedia: <https://arxiv.org/abs/1404.1100>
- [3] Munir, R., "Nilai Eigen dan Vektor Eigen (Bagian 1)," Institut Teknologi Bandung, Okt. 2025. [Daring]. Tersedia: <https://informatika.stei.itb.ac.id/~rinaldi.munir/AljabarGeometri/2025-2026/Algeo-19-Nilai-Eigen-dan-Vektor-Eigen-Bagian1-2025.pdf>
- [4] Munir, R., "Singular Value Decomposition (SVD) (Bagian 1)," Institut Teknologi Bandung, Nov. 2025. [Daring]. Tersedia: <https://informatika.stei.itb.ac.id/~rinaldi.munir/AljabarGeometri/2025-2026/Algeo-21-Singular-value-decomposition-Bagian1-2025.pdf>

7 LAMPIRAN

1. Tautan menuju repositori GitHub kelompok: <https://github.com/IRK-23/algeo2-lemonnipis>
2. Tautan menuju hasil eksperimen versi lengkap: <https://drive.google.com/drive/folders/1WEPfYjPHzVM2SNuR5K9DJJ9CSh3j4MAI?usp=sharing>