

**SISTEM REKOMENDASI FILM MENGGUNAKAN METODE
CONTENT-BASED FILTERING BERDASARKAN SINOPSIS FILM**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh

ILHAM PUTRA FIRMANSYAH

21.11.3868

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**SISTEM REKOMENDASI FILM MENGGUNAKAN METODE
CONTENT-BASED FILTERING BERDASARKAN SINOPSIS FILM**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh

ILHAM PUTRA FIRMANSYAH

21.11.3868

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

SISTEM REKOMENDASI FILM MENGGUNAKAN METODE *CONTENT-BASED FILTERING* BERDASARKAN SINOPSIS FILM

yang disusun dan diajukan oleh

Ilham Putra Firmansyah

21.11.3868

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 18 Februari 2025

Dosen Pembimbing,



Anna Baita S.Kom., M.Kom.

NIK. 190302290

HALAMAN PENGESAHAN
SKRIPSI
**SISTEM REKOMENDASI FILM MENGGUNAKAN METODE *CONTENT-
BASED FILTERING* BERDASARKAN SINOPSIS FILM**

yang disusun dan diajukan oleh

Ilham Putra Firmansyah

21.11.3868

Telah dipertahankan di depan Dewan Penguji
pada tanggal 18 Februari 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Ferian Fauzi Abdulloh, S.Kom., M.Kom.
NIK. 190302276

Ike Verawati, S.Kom., M.Kom.
NIK. 190302237

Anna Baita, S.Kom., M.Kom.
NIK. 190302290

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 18 Februari 2025

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Ilham Putra Firmansyah
NIM : 21.11.3868

Menyatakan bahwa Skripsi dengan judul berikut:

SISTEM REKOMENDASI FILM MENGGUNAKAN METODE *CONTENT-BASED FILTERING* BERDASARKAN SINOPSIS FILM

Dosen Pembimbing : Anna Baita, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar **ASLI** dan **BELUM PERNAH** diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian **SAYA** sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab **SAYA**, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini **SAYA** buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka **SAYA** bersedia menerima **SANKSI AKADEMIK** dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 18 Februari 2025

Yang Menyatakan,

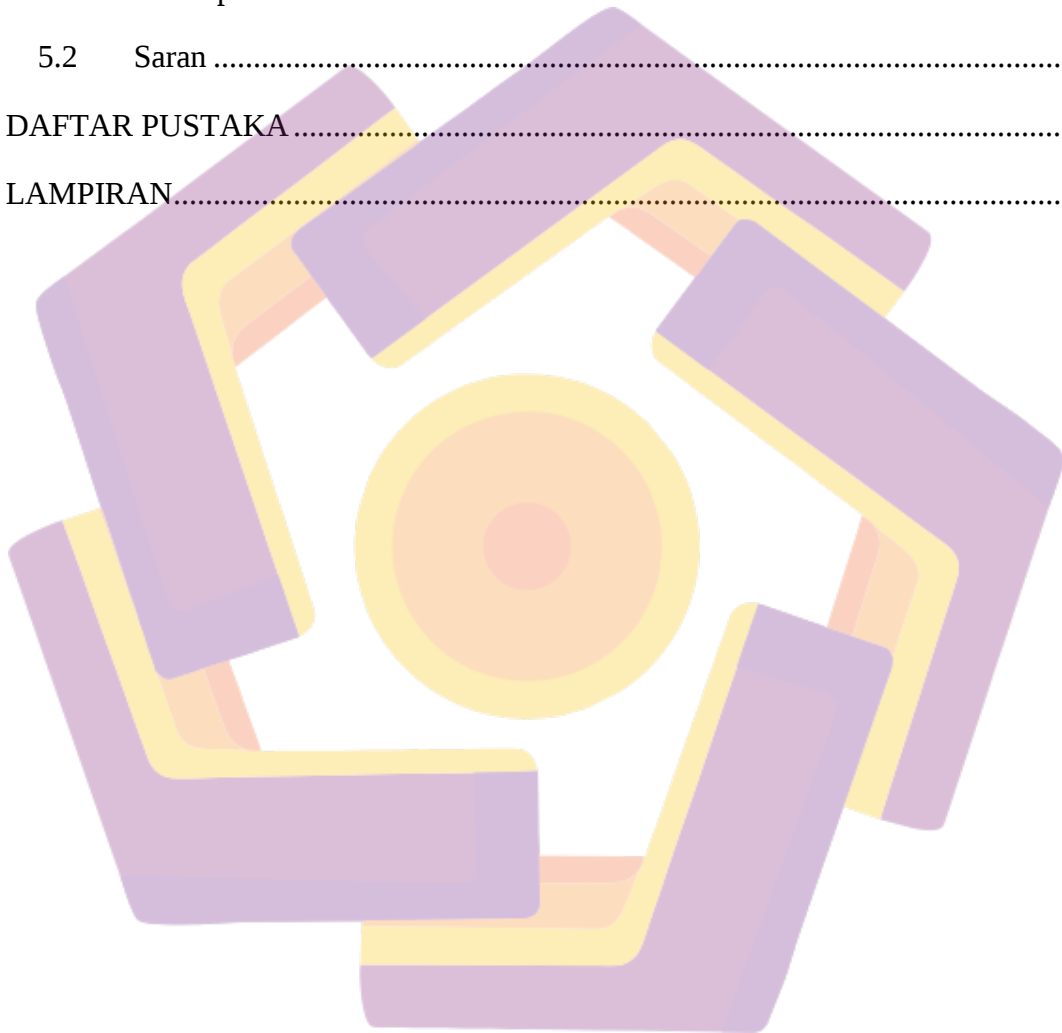


Ilham Putra Firmansyah

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
DAFTAR ISI.....	v
DAFTAR TABEL.....	vii
DAFTAR GAMBAR	viii
DAFTAR LAMPIRAN.....	ix
INTISARI	x
ABSTRACT.....	xi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.2 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori	14
BAB III METODE PENELITIAN	21
3.1 Objek Penelitian.....	21
3.2 Alur Penelitian	21
3.3 Alat dan Bahan.....	32
BAB IV HASIL DAN PEMBAHASAN	34

4.1	Profile Data	34
4.2	Akuisisi Data.....	34
4.3	Evaluasi.....	36
4.4	<i>Deployment</i>	42
BAB V PENUTUP		45
5.1	Kesimpulan	45
5.2	Saran	45
DAFTAR PUSTAKA		46
LAMPIRAN.....		49

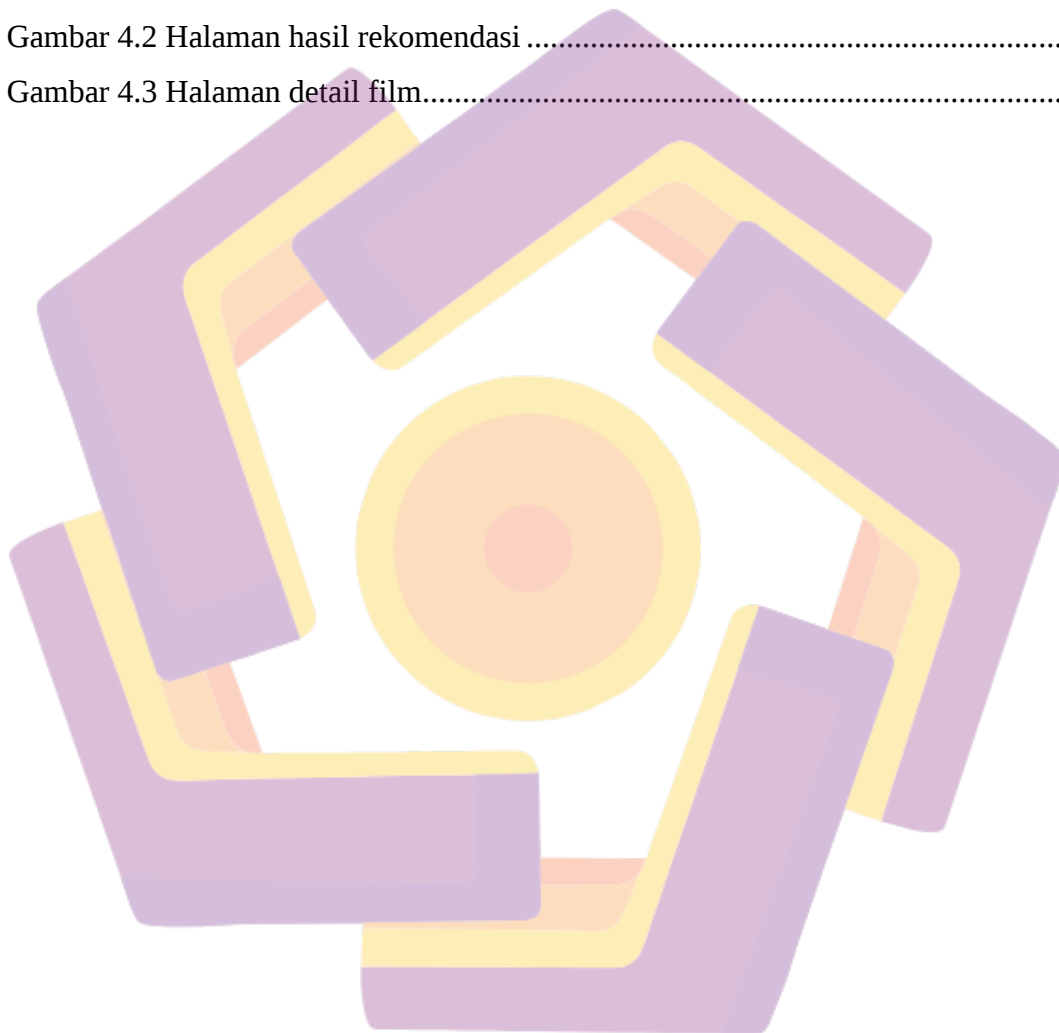


DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	7
Tabel 3.1 Dataset Film.....	22
Tabel 3.2 Contoh Data Stopword Removal	24
Tabel 3.3 Deskripsi Data Overview	26
Tabel 3.4 Frekuensi data film per tahun rilis	27
Tabel 3.5 Frekuensi Kata	28
Tabel 3.6 Hasil TF-IDF pada film pertama	29
Tabel 3.7 Contoh Data Numerik Sinopsis Film.....	30
Tabel 4.1 Dataset Film setelah slicing	35
Tabel 4.2 Hasil rekomendasi dan ground truth cosine similarity	37
Tabel 4.3 Hasil precision dan recall cosine similarity	38
Tabel 4.4 Hasil rekomendasi dan ground truth euclidean distances	39
Tabel 4.5 Hasil precision dan recall euclidean distances.....	40
Tabel 4.6 Perbandingan cosine similarity dan euclidean distances	41

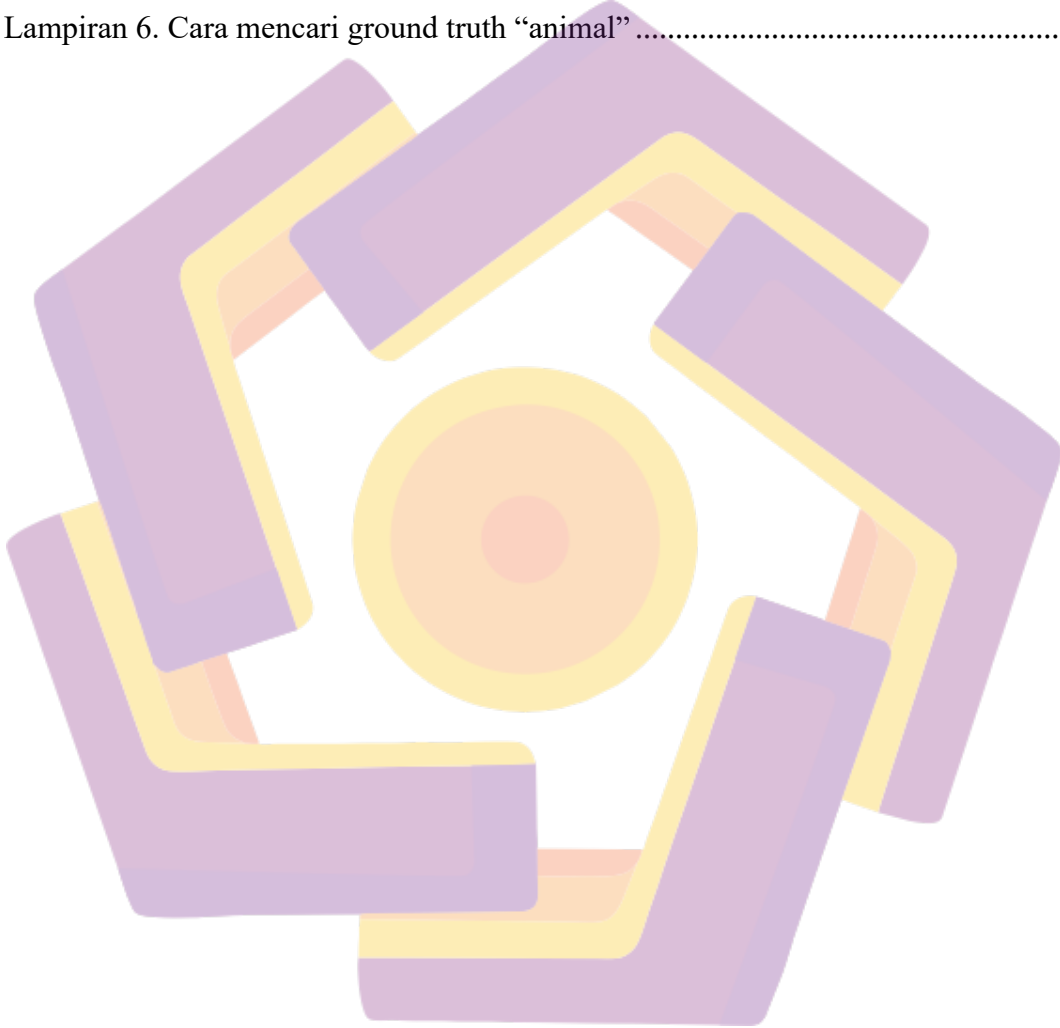
DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian	21
Gambar 3.2 Distribusi Jumlah Kata.....	25
Gambar 3.3 Distribusi Tahun Rilis Film	27
Gambar 3.4 Visualisasi <i>Wordcloud</i>	29
Gambar 4.1 Halaman utama	42
Gambar 4.2 Halaman hasil rekomendasi	43
Gambar 4.3 Halaman detail film.....	44



DAFTAR LAMPIRAN

Lampiran 1. Kode pembuatan model cosine similarity	49
Lampiran 2. Kode pembuatan model euclidean distances.....	54
Lampiran 3. Kode deployment sistem	59
Lampiran 4. Cara mencari ground truth “spider”	61
Lampiran 5. Cara mencari ground truth “toy”	65
Lampiran 6. Cara mencari ground truth “animal”	69



INTISARI

Banyaknya film yang telah diproduksi hingga saat ini membuat pencarian film secara manual tentu akan memakan banyak waktu, menonton film di bioskop, platform online penyedia streaming film, maupun membeli kaset DVD juga akan memakan biaya jika film yang ditonton tidak sesuai keinginan. Solusi yang dapat mengatasi permasalahan sulitnya mencari film yang diinginkan adalah dengan membuat suatu sistem rekomendasi film. Sistem rekomendasi film dapat dibangun menggunakan beberapa metode. Pada penelitian ini metode yang digunakan untuk membangun sistem rekomendasi film adalah *content based filtering* dengan memanfaatkan fitur *overview* atau sinopsis film. Pencarian kemiripan antar data film menggunakan *cosine similarity* dan *euclidean distances*. Dari kedua metode tersebut kemudian dibandingkan hasil rekomendasi film yang dihasilkan. Dari tiga *keyword* berbeda yang telah dilakukan ujicoba yaitu “*spider*”, “*toy*”, dan “*animal*” didapatkan rata-rata nilai *precision* dan *recall* masing-masing 0,9000 dan 1,0000. Hal ini membuktikan bahwa sistem rekomendasi film menggunakan metode *content based filtering* berdasarkan sinopsis film menghasilkan rekomendasi film yang relevan. Tidak terdapat perbedaan antara metode *cosine similarity* dengan *euclidean distances* dalam menentukan kemiripan data film sebagai hasil rekomendasi. Dengan sistem rekomendasi film yang dihasilkan, para penikmat film dapat dengan mudah menemukan rekomendasi film yang tepat untuk ditonton.

Kata kunci: *Content Based Filtering, Cosine Similarity, Euclidean Distances, Film, Sinopsis.*

ABSTRACT

The large number of films that have been produced to date means that searching for films manually will certainly take a lot of time, watching films in theaters, online platforms that provide film streaming, or buying DVD cassettes will also cost money if the film watched is not as desired. A solution that can overcome the problem of difficulty in finding the desired film is to create a film recommendation system. A film recommendation system can be built using several methods. In this study, the method used to build a film recommendation system is content-based filtering by utilizing the overview or film synopsis feature. Searching for similarities between film data using cosine similarity and euclidean distances. The results of the resulting film recommendations are then compared from the two methods. From the three different keywords that have been tested, namely "spider", "toy", and "animal", the average precision and recall values were 0.9000 and 1.0000, respectively. This proves that the film recommendation system using the content-based filtering method based on the film synopsis produces relevant film recommendations. There is no difference between the cosine similarity and euclidean distances methods in determining the similarity of film data as a recommendation result. With the resulting movie recommendation system, movie lovers can easily find the right movie recommendations to watch.

Keyword: Content Based Filtering, Cosine Similarity, Euclidean Distances, Movies, Synopsis.

