

# LAPORAN TUGAS AKHIR

ANALISIS SENTIMEN DIALOG FILM MENGGUNAKAN  
COMPLEMENT NAÏVE BAYES



**DISUSUN OLEH :**

**SALMA FARADHIBA**

**NIM: 04221032**

**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS NAROTAMA  
2025**

# LAPORAN TUGAS AKHIR

ANALISIS SENTIMEN DIALOG FILM MENGGUNAKAN  
COMPLEMENT NAÏVE BAYES



**DISUSUN OLEH :**

**SALMA FARADHIBA**

**NIM: 04221032**

**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS NAROTAMA  
2025**

LEMBAR PENGESAHAN

LAPORAN TUGAS AKHIR

ANALISIS SENTIMEN DIALOG FILM  
MENGUNAKAN COMPLEMENT NAÏVE BAYES

DISUSUN OLEH:

SALMA FARADHIBA

NIM: 04221032

Dipertahankan di depan Penguji Ujian Tugas Akhir  
Program Studi Sistem Informasi  
Fakultas Ilmu Komputer  
Universitas Narotama Surabaya  
Tanggal 24 Juli 2025

Penguji



1. Eman Setiawan, S.Kom., M.M.  
NIDN: 0720017501

Ketua Program Studi



Eman Setiawan, S.Kom., M.M.  
NIDN: 0720017501

2. Dr. Arvo Nugroho, S.T., S.Kom., M.T.  
NIDN: 0721077001



3. Awalludivah Ambarwati, S.Kom., M.M.  
NIDN: 0015047801

Fakultas Ilmu Komputer



Dr. Cahyo Daryajati, S.T., M.T.  
NIDN: 0710097402



LEMBAR PERSETUJUAN


## LAPORAN TUGAS AKHIR

ANALISIS SENTIMEN DIALOG FILM  
MENGUNAKAN COMPLEMENT NAÏVE BAYES

DISUSUN OLEH:

SALMA FARADHIBA

NIM: 04221032



Diajukan guna memenuhi persyaratan  
Untuk memperoleh gelar Sarjana Komputer (S.Kom)  
Pada Program Studi Sistem Informasi  
Fakultas Ilmu Komputer  
Universitas Narotama Surabaya

Surabaya, 24 Juli 2025

Mengetahui/Menyetujui

Dosen Pembimbing,



Dr. Aryo Nugroho, S.T., S.Kom., M.T

NIDN: 0721077001

## SURAT PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Laporan Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan disuatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat Karya/Pendapat yang pernah ditulis oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Acuan/Daftar Pustaka.

Apabila ditemukan suatu Jiplakan/Plagiat maka saya bersedia menerima akibat berupa sanksi Akademis dan sanksi lain yang diberikan oleh yang berwenang sesuai ketentuan peraturan dan perundang-undangan yang berlaku.



## KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT atas segala limpahan rahmat, hidayah, dan karunia-Nya yang telah memberikan kesehatan, kekuatan, dan kesempatan sehingga penulis dapat menyelesaikan skripsi ini. Shalawat serta salam semoga senantiasa tercurah kepada junjungan Nabi Muhammad SAW, suri teladan sepanjang masa, yang ajarannya senantiasa menjadi cahaya dalam perjalanan hidup. Skripsi dengan judul “Analisis Sentimen Dialog Film Menggunakan Complement Naive Bayes” ini disusun untuk memenuhi salah satu syarat memperoleh gelar Sarjana pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Narotama Surabaya.

Penyusunan skripsi ini tentu tidak akan terselesaikan tanpa doa, dukungan, bimbingan, dan semangat dari berbagai pihak. Oleh karena itu, dengan penuh rasa hormat dan tulusnya hati, penulis menyampaikan terima kasih kepada:

1. Allah SWT yang telah melimpahkan rahmat-Nya sehingga diberikan kesempatan, kemudahan dan kelancaran dalam melaksanakan penelitian hingga penyusunan skripsi ini.
2. Bapak Dr. Arasy Alimudin, SE, MM selaku Rektor Universitas Narotama Surabaya.
3. Bapak Eman Setiawan, S. Kom., M. M. selaku Ketua Program Studi Sistem Informasi.
4. Bapak Dr. Aryo Nugroho, S. T., S. Kom., M. T. selaku Dosen Pembimbing serta penguji yang telah memberikan arahan dan inspirasi mulai dari penelitian hingga penyusunan skripsi.

5. Seluruh dosen Fakultas Ilmu Komputer Universitas Narotama Surabaya yang telah membekali penulis dengan berbagai ilmu selama mengikuti perkuliahan sampai akhir penulisan skripsi.
6. Almarhum kedua orang tua yang telah berpulang, doa dan kasih sayang kalian adalah kekuatan abadi yang tak pernah pudar, menjadi alasan penulis terus melangkah hingga titik ini.
7. Ayah tiri dan adik, yang dengan ketulusan hati memberikan kasih sayang, perhatian, serta menjadi sumber kehangatan di setiap langkah penulis.
8. Mama (kakak ibu) dan adik-adik ibu, yang senantiasa menghadirkan rasa rumah, memberi pelukan di saat penulis lelah, serta menjadi pengingat bahwa keluarga adalah tempat terbaik untuk kembali.
9. Teman seperjuangan Program Studi Sistem Informasi Angkatan 2021 yang telah menjadi bagian dari cerita indah perjuangan ini, berbagi tawa, lelah, dan semangat tanpa henti.
10. Terakhir, untuk diri sendiri. Terima kasih telah memilih untuk tidak menyerah ketika badai datang, tetap berdiri meski langkah terasa berat, dan terus berusaha menyelesaikan apa yang sudah dimulai. Terima kasih telah bertahan dalam malam-malam panjang yang penuh lelah dan ragu, bangkit di pagi-pagi yang kadang tak memberi kepastian, dan percaya bahwa setiap tetes keringat, setiap doa yang terucap, akan berbuah manis pada waktunya. Terima kasih telah belajar bahwa keberhasilan bukan hanya tentang hasil akhir, tetapi tentang perjalanan, proses, dan keberanian untuk terus melangkah walau hati pernah ingin berhenti.

Akhir kata penulis mengucapkan maaf yang sebesar-besarnya apabila terdapat kekurangan dan kesalahan baik dalam penulisan maupun penyusunan skripsi. Penulis menyadari bahwa terdapat kekurangan dalam penulisan skripsi ini, penulis mengharapkan kritik dan saran yang sifatnya membangun dari para pembaca. Penulis juga berharap skripsi ini dapat bermanfaat bagi pembaca.

Surabaya, 14 Agustus 2025

**Salma Faradhiba**

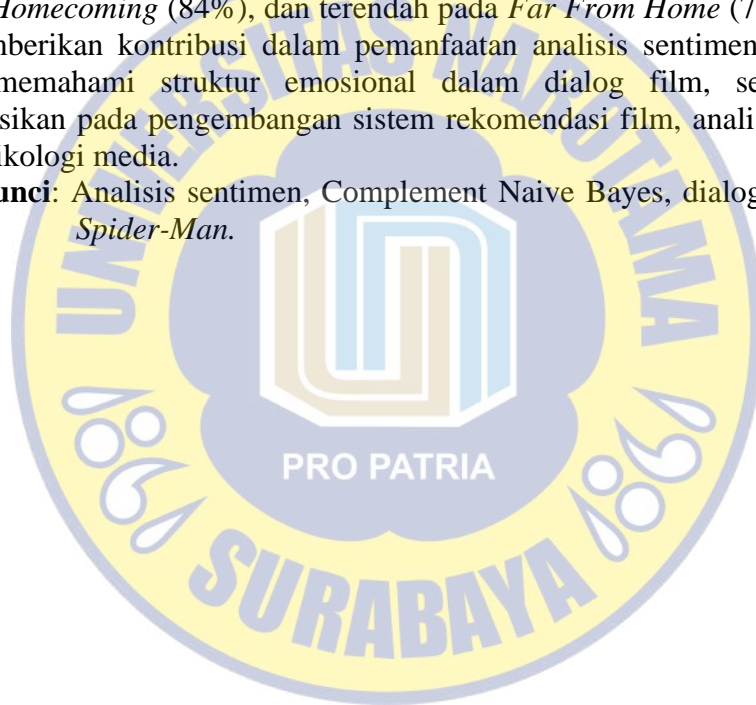


## ABSTRAK

Pesatnya perkembangan industri film serta meningkatnya minat masyarakat terhadap narasi yang emosional dan tematik menjadikan analisis dialog film sebagai salah satu bidang kajian yang relevan dalam pemrosesan bahasa alami. Penelitian ini bertujuan untuk mengklasifikasikan sentimen dalam dialog film *Spider-Man* ke dalam tiga kategori tematik, yaitu Konflik, Persahabatan, dan *Multiverse*, dengan menggunakan metode Complement Naive Bayes (CNB). Data diambil dari transkripsi dialog tiga film *Spider-Man* versi Tom Holland: *Homecoming* (2017), *Far From Home* (2019), dan *No Way Home* (2021), yang kemudian diproses melalui tahapan *text cleaning*, *tokenizing*, *stopword removal*, dan *stemming*.

Hasil penelitian menunjukkan bahwa model CNB mampu melakukan klasifikasi sentimen tematik dengan tingkat akurasi tertinggi pada film *No Way Home* (94%), diikuti *Homecoming* (84%), dan terendah pada *Far From Home* (77%). Penelitian ini memberikan kontribusi dalam pemanfaatan analisis sentimen berbasis tema untuk memahami struktur emosional dalam dialog film, serta berpotensi diaplikasikan pada pengembangan sistem rekomendasi film, analisis naskah, dan studi psikologi media.

**Kata kunci:** Analisis sentimen, Complement Naive Bayes, dialog film, TF-IDF, *Spider-Man*.

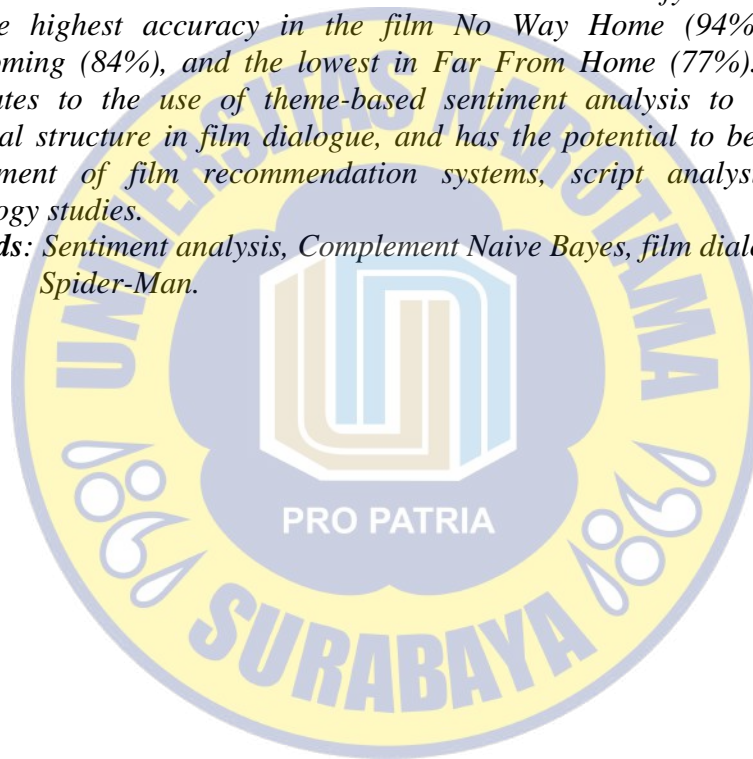


## ABSTRACT

*The rapid development of the film industry and the increasing public interest in emotional and thematic narratives make film dialogue analysis one of the relevant fields of study in natural language processing. This study aims to classify the sentiments in the dialogue of the Spider-Man film into three thematic categories, namely Conflict, Friendship, and Multiverse, using the Complement Naive Bayes (CNB) method. Data was taken from the dialogue transcriptions of three Tom Holland versions of Spider-Man: Homecoming (2017), Far From Home (2019), and No Way Home (2021), which were then processed through the stages of text cleaning, tokenizing, stopwords removal, and stemming.*

*The results showed that the CNB model was able to classify thematic sentiments with the highest accuracy in the film No Way Home (94%), followed by Homecoming (84%), and the lowest in Far From Home (77%). This research contributes to the use of theme-based sentiment analysis to understand the emotional structure in film dialogue, and has the potential to be applied to the development of film recommendation systems, script analysis, and media psychology studies.*

**Keywords:** *Sentiment analysis, Complement Naive Bayes, film dialogues, TF-IDF, Spider-Man.*



## DAFTAR ISI

LEMBAR PERSETUJUAN.....	iv
LEMBAR PENGESAHAN .....	v
SURAT PERNYATAAN.....	vi
MOTTO DAN PERSEMBAHAN .....	vii
KATA PENGANTAR .....	viii
ABSTRAK .....	xi
ABSTRACT.....	xii
DAFTAR ISI.....	xiii
DAFTAR GAMBAR .....	xvi
DAFTAR TABEL.....	xvii
DAFTAR LAMPIRAN.....	xviii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang .....	1
1.2. Rumusan Masalah .....	5
1.3. Batasan Masalah.....	5
1.4. Tujuan Penelitian.....	6
1.5. Manfaat Penelitian.....	6
BAB II TINJAUAN PUSTAKA.....	8
2.1 Teori – Teori Dasar .....	8
2.2.1. NLP ( <i>Natural Language Processing</i> ) .....	8
2.2.2. Text Mining.....	10
2.2.3. Analisis Sentimen .....	11
2.2.4. Dialog Film .....	13

2.2.5.	Teori Himpunan .....	14
2.2.6.	Teori Probabilitas .....	15
2.2.7.	Naïve Bayes .....	16
2.2.8.	Complement Naïve Bayes.....	17
2.2	Tinjauan Penelitian Terdahulu .....	18
BAB III METODOLOGI PENELITIAN.....		22
3.1.	Jenis Penelitian .....	22
3.2.	Desain Penelitian .....	22
3.3.	Sumber Data .....	31
3.4.	Teknik Pengumpulan Data .....	32
3.5.	Instrumen Penelitian.....	33
3.6.	Analisis Data .....	38
BAB IV HASIL DAN PEMBAHASAN .....		40
4.1	Hasil Pengambilan Data .....	40
4.2	Hasil Pengolahan Data ( <i>Preprocessing</i> ).....	42
4.3	Pemodelan .....	45
4.4	Evaluasi .....	45
4.4.1	<i>Term Frequency-Inverst Document Frequency</i> .....	46
4.4.2	Evaluasi Model: Confusion Matrix.....	51
4.4.3	Akurasi .....	60
4.5	Visualisasi Wordcloud .....	63
4.6	Pembahasan .....	72
BAB V KESIMPULAN DAN SARAN.....		75
5.1	Kesimpulan.....	75
5.2	Saran .....	76

DAFTAR PUSTAKA ..... 78

LAMPIRAN ..... 81

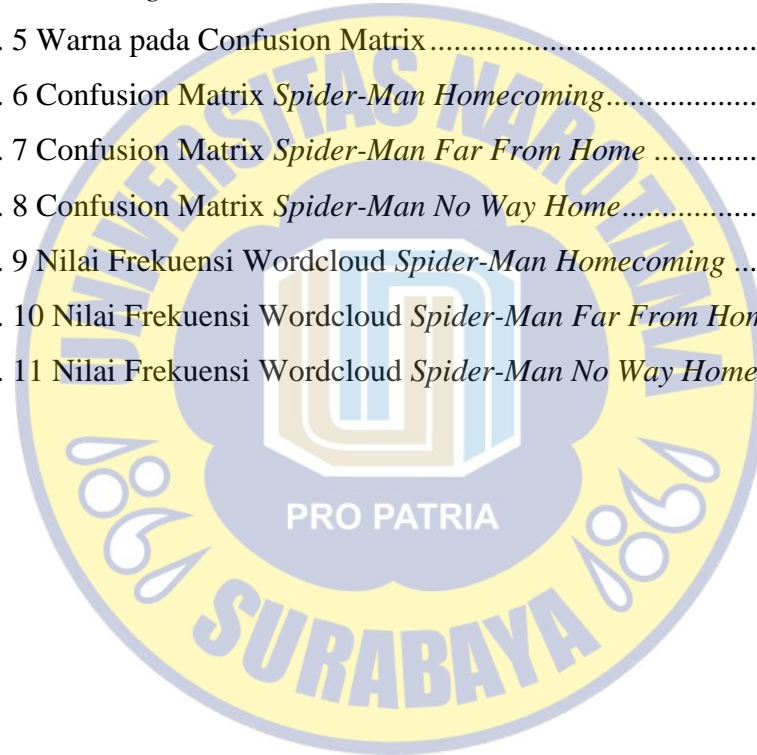


## DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian.....	23
Gambar 3. 2 <i>Source Code</i> Pengambilan Data.....	25
Gambar 3. 3 <i>Source code preprocessing</i> .....	26
Gambar 3. 4 <i>Source Code</i> Pemodelan .....	29
Gambar 4. 1 Hasil Transkripsi dari Audio .....	41
Gambar 4. 2 Dialog Film .....	41
Gambar 4. 3 <i>Source Code</i> TF-IDF.....	47
Gambar 4. 4 TF-IDF <i>Spider-Man Homecoming</i> .....	47
Gambar 4. 5 TF-IDF <i>Spider-Man Far From Home</i> .....	49
Gambar 4. 6 TF-IDF <i>Spider-Man No Way Home</i> .....	50
Gambar 4. 7 <i>Source Code</i> Confusion Matrix.....	52
Gambar 4. 8 Confusion Matrix <i>Spider-Man Homecoming</i> .....	55
Gambar 4. 9 Confusion Matrix <i>Spider-Man Far From Home</i> .....	58
Gambar 4. 10 Confusion Matrix <i>Spider-Man No Way Home</i> .....	60
Gambar 4. 11 <i>Source Code</i> Akurasi.....	62
Gambar 4. 12 Perbandingan Akurasi .....	62
Gambar 4. 13 <i>Source Code</i> Visualisasi Wordcloud.....	65
Gambar 4. 14 Wordcloud <i>Spider-Man Homecoming</i> .....	66
Gambar 4. 15 Wordcloud <i>Spider-Man Far From Home</i> .....	68
Gambar 4. 16 Wordcloud <i>Spider-Man No Way Home</i> .....	70

## DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu .....	19
Tabel 3. 1 Alat Bantu Ekstraksi dan Transkripsi .....	34
Tabel 3. 2 Pustaka Python untuk NLP dan Analisis .....	36
Tabel 4. 1 <i>Text Cleaning</i> .....	42
Tabel 4. 2 <i>Tokenization</i> .....	43
Tabel 4. 3 <i>Stopword</i> .....	44
Tabel 4. 4 <i>Stemming</i> .....	44
Tabel 4. 5 Warna pada Confusion Matrix .....	53
Tabel 4. 6 Confusion Matrix <i>Spider-Man Homecoming</i> .....	54
Tabel 4. 7 Confusion Matrix <i>Spider-Man Far From Home</i> .....	56
Tabel 4. 8 Confusion Matrix <i>Spider-Man No Way Home</i> .....	58
Tabel 4. 9 Nilai Frekuensi Wordcloud <i>Spider-Man Homecoming</i> .....	65
Tabel 4. 10 Nilai Frekuensi Wordcloud <i>Spider-Man Far From Home</i> .....	67
Tabel 4. 11 Nilai Frekuensi Wordcloud <i>Spider-Man No Way Home</i> .....	69



## DAFTAR LAMPIRAN

Lampiran 1 Pengajuan Proposal.....	81
Lampiran 2 Surat Pengantar Bimbingan .....	82
Lampiran 3 Daftar Bimbingan .....	83
Lampiran 4 Daftar Perbaikan .....	84
Lampiran 5 Jurnal .....	87
Lampiran 6 Bukti Submit Jurnal .....	95
Lampiran 7 Poster .....	96
Lampiran 8 Cek Plagiasi .....	97

