

**PERBANDINGAN ANALISIS SENTIMEN APLIKASI *CHATGPT* DAN *DEEPSEEK*
MENGUNAKAN METODE *NAIVE BAYES CLASSIFIER* DAN KONTEN
ANALISIS**

Tugas Akhir

Diajukan untuk memenuhi persyaratan mencapai derajat Sarjana Sistem Informasi



**Fransisca Mayang Sukma Adiana
NPM: 211711060**

**Program Studi Sistem Informasi
Fakultas Teknologi Industri
Universitas Atma Jaya Yogyakarta
2025**

LEMBAR PENGESAHAN

HALAMAN PENGESAHAN

Tugas Akhir Berjudul

PERBANDINGAN ANALISIS SENTIMEN APLIKASI CHATGPT DAN DEEPSEEK MENGGUNAKAN METODE
NAIVE BAYES CLASSIFIER DAN KONTEN ANALISIS

yang disusun oleh

Fransisca Mayang Sukma Adiana

211711060

dinyatakan telah memenuhi syarat pada tanggal 06 Juni 2025

		Keterangan
Dosen Pembimbing 1	: Generosa Lukhayu Pritalia, S.T., M.Eng.	Telah Menyetujui
Dosen Pembimbing 2	: Dr. FI. Spty Rahayu, S.T., M.Kom.	Telah Menyetujui
Tim Penguji		
Penguji 1	: Generosa Lukhayu Pritalia, S.T., M.Eng.	Telah Menyetujui
Penguji 2	: Emanuel Ristian Handoyo, S.T., M.Eng.	Telah Menyetujui
Penguji 3	: Elisabeth Marsella, S.S., M.Li.	Telah Menyetujui

Yogyakarta, 06 Juni 2025

Universitas Atma Jaya Yogyakarta

Teknologi Industri

Dekan

ttd.

Dr. Ir. Parama Kartika Dewa SP., S.T., M.T.

Dokumen ini merupakan dokumen resmi UAJY yang tidak memerlukan tanda tangan karena dihasilkan secara elektronik oleh Sistem Bimbingan UAJY. UAJY bertanggung jawab penuh atas informasi yang tertera di dalam dokumen ini

LEMBAR PERNYATAAN KEASLIAN PENELITIAN

Saya yang bertanda tangan di bawah ini:

Nama Lengkap : Fransisca Mayang Sukma Adiana

NPM : 211711060

Program Studi : Sistem Informasi

Fakultas : Fakultas Teknologi Industri

Judul Penelitian : PERBANDINGAN ANALISIS SENTIMEN APLIKASI *CHATGPT* DAN *DEEPSEEK* MENGGUNAKAN METODE *NAIVE BAYES CLASSIFIER* DAN KONTEN ANALISIS

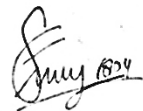
Menyatakan dengan ini:

1. Skripsi ini adalah benar merupakan hasil karya sendiri dan tidak merupakan salinan sebagian atau keseluruhan dari karya orang lain.
2. Memberikan kepada Universitas Atma Jaya Yogyakarta, berupa Hak Bebas Royalti non eksklusif (*Non-Exclusive-Royalty-Free Right*) atas Penelitian ini, dan berhak menyimpan, mengelola dalam pangkalan data, mendistribusikan, serta menampilkan untuk kepentingan akademis, tanpa perlu meminta izin selama tetap mencantumkan nama penulis.
3. Bersedia menanggung secara pribadi segala bentuk tuntutan hukum yang mengikuti atas pelanggaran Hak Cipta dalam pembuatan skripsi ini.

Demikianlah pernyataan ini dibuat dan dapat dipergunakan sebagaimana mestinya.

Yogyakarta, 2 Juni 2025

Yang menyatakan,



Fransisca Mayang Sukma Adiana

211711060

PRAKATA

Puji dan syukur kehadiran Tuhan yang Maha Esa atas rahmat dan berkatnya peneliti dapat menyelesaikan tugas akhir sebagai syarat penuntasan derajat Sarjana Sistem Informasi dengan penelitian berjudul "Perbandingan Analisis Sentimen Aplikasi *ChatGPT* dan *DeepSeek* Menggunakan Metode *Naive Bayes Classifier* dan Konten Analisis." Pada kesempatan ini, peneliti ingin mengucapkan terima kasih kepada berbagai pihak yang sudah mendukung, menemani, dan memberi bantuan kepada peneliti dalam proses pengerjaan tugas akhir ini, yaitu kepada :

1. Tuhan Yesus Kristus dan Bunda Maria atas berkat, petunjuk, dan kasih karunia yang diberikan selama peneliti mengerjakan tugas akhir ini.
2. Ibu Generosa Lukhayu Pritalia, S.T., M.Eng., selaku dosen pembimbing 1 dan dosen pembimbing akademik yang selalu sabar dalam memberikan pengarahan, mendukung, menyemangati, serta membimbing peneliti dalam menyelesaikan tugas akhir.
3. Ibu Dr. Florensia Spty Rahayu, S.T., M.Kom., selaku dosen pembimbing 2 yang mendukung , mengarahkan, dan membimbing selama tugas akhir ini.
4. Papa, Mama, Bapak, Ibu, Bude-Bude, dan keluarga atas doa, kasih sayang, semangat, motivasi, dan dukungan yang diberikan kepada peneliti dari peneliti lahir hingga bisa menyelesaikan tugas akhir.
5. Kelompok "Emang Boleh Seteam Ini?", yaitu Brigitta Christa N. A, Dylan Saputra. W, Yohanes Kevin. W. U, Gregorius Yogiswara. V. W, Felix Wilson, Mario Steven, Malvin Christy, Kevin Surya, dan Chandra Wijaya. K yang selalu kebersamai, saling mendukung, menyemangati, membagi pengetahuan dengan peneliti selama perkuliahan dari awal hingga akhir.
6. Teman-teman SS KKP, Cirebon, HIMSI, FTI, SI, UAJY, dan lainnya yang tidak bisa peneliti sebutkan satu-satu, atas dukungan dalam setiap proses pendidikan dan kehidupan yang peneliti jalani hingga akhirnya menyelesaikan pendidikan ini.

Semoga penelitian tugas akhir ini dapat berkontribusi positif dalam bidang keilmuan sistem informasi, khususnya mengenai *machine learning* dan menjadi referensi untuk penelitian dan pengembangan pendidikan berikutnya. Besar harapan peneliti, hasil tugas akhir ini dapat bermanfaat bagi semua pihak yang membutuhkan.

ABSTRAK

Pada era kemajuan dan perkembangan teknologi yang semakin masif ke depan. Teknologi informasi terus berkembang dengan adanya *Artificial Intelligence*. Penelitian ini bertujuan untuk membandingkan analisis sentimen dari dua aplikasi *generative AI*, yaitu *ChatGPT* dan *DeepSeek*, berdasarkan ulasan pengguna di *Google Playstore*. Penelitian menggunakan metode *Naive Bayes Classifier* (khususnya *Multinomial Naive Bayes*) untuk analisis sentimen dan konten analisis menggunakan perangkat lunak *QDA Miner*. Data ulasan dikumpulkan melalui proses *scraping* pada periode Januari hingga April 2025, lalu dilakukan proses *pre-processing* seperti *case folding*, *normalisasi*, *tokenization*, *stopword removal*, dan *lemmatization*. Hasil analisis menunjukkan bahwa *ChatGPT* memperoleh proporsi sentimen positif lebih tinggi, yaitu 30.744 ulasan dibandingkan *DeepSeek* hanya 1.787 ulasan dengan angka akurasi *ChatGPT* 95% dan *DeepSeek* 84%, yang menunjukkan bahwa pengguna cenderung memiliki pengalaman yang lebih memuaskan terhadap *ChatGPT*. Selain itu, ditemukan pula beberapa faktor dominan yang memengaruhi persepsi pengguna seperti performa aplikasi, biaya, kecepatan respons, dan kualitas jawaban. Sementara untuk rating, *ChatGPT* lebih banyak memiliki rating diatas 4.4 dan stabil dibandingkan *DeepSeek*.

Kata Kunci : Analisis Sentimen, Konten Analisis, *ChatGPT*, *DeepSeek*, *Naive Bayes Classifier*, *Multinomial Naive Bayes*, *Google Playstore*, *Machine Learning*.

ABSTRACT

In an era of rapid technological advancement, information technology continues to evolve with the emergence of Artificial Intelligence. This study aims to compare sentiment analysis of two generative AI applications, ChatGPT and DeepSeek based on user reviews from the Google Play Store. The Naive Bayes Classifier method (specifically, the Multinomial Naive Bayes) was used for sentiment analysis, along with content analysis conducted using QDA Miner software. Review data was collected through a scraping process from January to April 2025, followed by pre-processing steps including case folding, normalization, tokenization, stopword removal, and lemmatization. The analysis results showed that ChatGPT received a higher proportion of positive sentiment, with 30,744 positive reviews compared to only 1,787 for DeepSeek. The sentiment classification accuracy was also higher for ChatGPT at 95%, compared to 84% for DeepSeek, indicating that users tend to have a more satisfying experience with ChatGPT. Additionally, several dominant factors influencing user perception were identified, such as application performance, cost, response speed, and the quality of answers. In terms of ratings, ChatGPT consistently received ratings above 4.4 and demonstrated greater stability compared to DeepSeek.

Keywords : Sentiment Analysis, Content Analysis, ChatGPT, DeepSeek, Naive Bayes Classifier, Multinomial Naive Bayes, Google Playstore, Machine Learning.

DAFTAR ISI

PRAKATA	IV
ABSTRAK.....	V
ABSTRACT	VI
DAFTAR ISI	VII
DAFTAR GAMBAR.....	XI
DAFTAR TABEL	XIII
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Pertanyaan Penelitian	3
1.4 Batasan Penelitian.....	3
1.5 Tujuan Penelitian.....	3
1.6 Manfaat Penelitian.....	4
1.7 Bagan Keterkaitan	5
Bab II TINJAUAN PUSTAKA	6
2.1 Studi Sebelumnya	6
2.2 Dasar Teori	11
2.2.1 Analisis Sentimen	11
2.2.2 Modul <i>Generative AI Chatbot</i>	11
2.2.3 <i>Classifier</i>	12
2.2.4 <i>Confusion Matrix</i>	13
2.2.5 Konten Analisis.....	14
2.3 Tabel Perbandingan.....	14
BAB III METODE PENELITIAN.....	15

3.1	Metode Penelitian	15
3.2	Tahapan Penelitian.....	16
3.2.1	<i>Text Mining</i>	17
3.2.2	<i>Pre-Processing</i>	17
3.2.3	<i>Sentiment Analysis Naïve Bayes Classifier</i>	19
3.2.4	<i>Content Analysis</i>	21
3.2.5	<i>Trend Analysis</i>	22
3.3	<i>Tools yang Digunakan</i>	23
BAB IV HASIL dan PEMBAHASAN ANALISIS SENTIMEN		24
4.1	Hasil.....	24
4.1.1	<i>Text Mining</i>	24
4.1.1.1	<i>Data Scraping</i>	24
4.1.1.2	<i>Translate</i>	30
4.1.1.3	<i>Cleaning</i>	32
4.1.2	<i>Pre-Processing</i>	34
4.1.2.1	<i>Case Folding</i>	34
4.1.2.2	<i>Normalization</i>	35
4.1.2.3	<i>Tokenization</i>	37
4.1.2.4	<i>Stopword Removal</i>	38
4.1.2.5	<i>Lemmatization</i>	39
4.1.2.6	<i>WordCloud and Word Frequency Before Labeling</i>	41
4.1.2.6.1	<i>ChatGPT</i>	43
4.1.2.6.2	<i>DeepSeek</i>	44
4.1.3	<i>Sentiment Analysis Naïve Bayes Classifier</i>	45
4.1.3.1	<i>Sentiment Labeling</i>	45
4.1.3.2	<i>Sentiment Frequency Calculation</i>	47

4.1.3.2.1	<i>ChatGPT</i>	48
4.1.3.2.2	<i>DeepSeek</i>	49
4.1.3.3	<i>Data Training and Testing</i>	49
4.1.3.3.1	<i>ChatGPT</i>	51
4.1.3.3.2	<i>DeepSeek</i>	52
4.1.3.4	<i>Model Evaluation and Optimization</i>	53
4.1.3.4.1	<i>ChatGPT</i>	57
4.1.3.4.2	<i>DeepSeek</i>	62
4.1.3.5	<i>Data Visualization</i>	68
4.1.3.5.1	<i>ChatGPT</i>	68
4.1.3.5.2	<i>DeepSeek</i>	70
4.2	Pembahasan.....	73
4.2.1	<i>Sentiment Analysis</i>	73
4.2.1.1	<i>ChatGPT</i>	73
4.2.1.2	<i>DeepSeek</i>	77
BAB V HASIL DAN PEMBAHASAN ANALISIS KONTEN.....		81
5.1	Hasil.....	81
5.1.1	<i>Forming Sub Categories</i>	81
5.1.2	<i>Forming Categories</i>	82
5.1.3	<i>Result and Discussion</i>	83
5.1.3.1	<i>ChatGPT</i>	83
5.1.3.2	<i>DeepSeek</i>	83
5.1.4	<i>WordCloud Visualization</i>	84
5.1.4.1	<i>ChatGPT</i>	84
5.1.4.2	<i>DeepSeek</i>	85
5.2	Pembahasan.....	85

5.2.1	<i>Forming Sub Categories</i>	85
5.2.2	<i>Forming Categories</i>	87
5.2.3	<i>Result and Discussion</i>	88
5.2.3.1	<i>ChatGPT</i>	88
5.2.3.2	<i>DeepSeek</i>	91
5.2.4	<i>WordCloud Visualization</i>	96
5.2.4.1	<i>ChatGPT</i>	96
5.2.4.2	<i>DeepSeek</i>	97
BAB VI HASIL DAN PEMBAHASAN ANALISIS TREN		98
6.1	Hasil.....	98
6.1.1	<i>Trend Analysis</i>	98
6.1.1.1	<i>Rating</i>	98
6.1.1.2	<i>Number of Reviews</i>	99
6.1.1.3	<i>Number of Users</i>	100
6.2	Pembahasan.....	100
6.2.1	<i>Trend Analysis of ChatGPT and DeepSeek</i>	100
6.3	Kelebihan dan Kekurangan	103
6.3.1	Kelebihan	103
6.3.2	Kekurangan	104
BAB VII KESIMPULAN DAN SARAN		105
7.1	Kesimpulan	105
7.2	Saran	106
REFERENSI		108
LAMPIRAN		112

DAFTAR GAMBAR

Gambar 1. 1. Bagan Keterkaitan	5
Gambar 3. 1. Bagan Tahapan Penelitian.....	16
Gambar 4. 1. Code Pengunduhan Scraping	26
Gambar 4. 2. Bagian 1 Code Scraping ChatGPT	26
Gambar 4. 3. Bagian 2 Code Scraping ChatGPT	27
Gambar 4. 4. Bagian 3 Code Scraping ChatGPT	27
Gambar 4. 5. Jumlah Ulasan ChatGPT Scraping Data.....	28
Gambar 4. 6. Bagian 1 Code Scraping DeepSeek	28
Gambar 4. 7. Bagian 2 Code Scraping DeepSeek	29
Gambar 4. 8. Bagian 3 Code Scraping DeepSeek	29
Gambar 4. 9. Jumlah Ulasan Scraping Data DeepSeek	30
Gambar 4. 10. Kegagalan Translate	31
Gambar 4. 11. Code Hapus Duplikat.....	32
Gambar 4. 12. Bagian 1 Code Cleaning.....	33
Gambar 4. 13. Bagian 2 Code Cleaning.....	34
Gambar 4. 14. Tampilan Data Hasil Cleaning	34
Gambar 4. 15. Code Case Folding	35
Gambar 4. 16. Hasil Case Folding.....	35
Gambar 4. 17. Bagian 1 Code Normalisasi	36
Gambar 4. 18. Bagian 2 Code Normalisasi	36
Gambar 4. 19. Hasil Normalisasi	36
Gambar 4. 20. Code Tokenization	37
Gambar 4. 21. Hasil Tokenisasi	37
Gambar 4. 22. Code Stopword Removal.....	38
Gambar 4. 23. Hasil Stopword Removal	38
Gambar 4. 24. Code Lemmatization	40
Gambar 4. 25. Hasil Lemmatization	41
Gambar 4. 26. Code Wordcloud	42
Gambar 4. 27. Code Frekuensi Kata	42
Gambar 4. 28. WordCloud ChatGPT Sebelum Labeling	43
Gambar 4. 29. Frekuensi Kata ChatGPT Sebelum Labeling	43
Gambar 4. 30. WordCloud DeepSeek Sebelum Labeling	44
Gambar 4. 31. Frekuensi Kata DeepSeek Sebelum Labeling.....	44
Gambar 4. 32. Bagian 1 Code Pelabelan Sentimen	45
Gambar 4. 33. Bagian 2 Code Skoring.....	46
Gambar 4. 34. Hasil Labeling dan Skoring DeepSeek	46
Gambar 4. 35. Hasil Labeling dan Skoring ChatGPT	47
Gambar 4. 36. Code Penghitungan Jumlah Sentimen	47
Gambar 4. 37. Jumlah Ulasan ChatGPT Berdasarkan Sentimen Sudah Labeling	48
Gambar 4. 38. Jumlah Ulasan DeepSeek Berdasarkan Sentimen Sudah Labeling.....	49
Gambar 4. 39. Code Data Training dan Testing	50
Gambar 4. 40. Code Distribusi Label Data Latih dan Tes.....	51
Gambar 4. 41. Jumlah Data Training dan Testing ChatGPT	51
Gambar 4. 42. Distribusi Data Label Training dan Testing ChatGPT	51
Gambar 4. 43. Jumlah Data Training dan Testing DeepSeek.....	52

Gambar 4. 44. Distribusi Data Label Training dan Testing DeepSeek	52
Gambar 4. 45. Bagian 1 Code Analisis Sentimen Multinomial NB	54
Gambar 4. 46. Bagian 2 Code Analisis Sentimen Multinomial NB	55
Gambar 4. 47. Bagian 3 Code Analisis Sentimen Multinomial NB	55
Gambar 4. 48. Distribusi Label ChatGPT Sebelum Resampling	57
Gambar 4. 49. Distribusi Label ChatGPT Setelah SMOTETomek.....	57
Gambar 4. 50. Hasil Analisis Sentimen Naïve Bayes Classifier ChatGPT	58
Gambar 4. 51. Confusion Matrix ChatGPT.....	60
Gambar 4. 52. Learning Curve ChatGPT	61
Gambar 4. 53. Distribusi Kasus Salah Klasifikasi ChatGPT	62
Gambar 4. 54. Distribusi Label DeepSeek Sebelum Resampling	62
Gambar 4. 55. Distribusi Label DeepSeek Setelah SMOTETomek	63
Gambar 4. 56. Hasil Analisis Sentimen Naïve Bayes Classifier DeepSeek	64
Gambar 4. 57. Confusion Matrix DeepSeek	65
Gambar 4. 58. Learning Curve DeepSeek	66
Gambar 4. 59. Distribusi Kasus Salah Klasifikasi DeepSeek.....	67
Gambar 4. 60. WordCloud Sentimen Positif ChatGPT.....	68
Gambar 4. 61. WordCloud Sentimen Negatif ChatGPT	69
Gambar 4. 62. Frekuensi Kata Sentimen Negatif dan Positif ChatGPT	69
Gambar 4. 63. WordCloud Sentimen Positif DeepSeek	70
Gambar 4. 64. WordCloud Sentimen Negatif DeepSeek.....	71
Gambar 4. 65. Frekuensi Kata Sentimen Negatif dan Positif DeepSeek.....	72
Gambar 5. 1. Pembentukan Sub Kategori	81
Gambar 5. 2. Sub Kategori	81
Gambar 5. 3. Pembentukan Kategori dan Sub Kategori.....	82
Gambar 5. 4. Kategori.....	82
Gambar 5. 5. Hubungan Antar Sub Kategori ChatGPT Berdasarkan QDA Miner	83
Gambar 5. 6. Hubungan Antar Sub Kategori DeepSeek Berdasarkan QDA Miner	83
Gambar 5. 7. WordCloud QDA Miner ChatGPT	84
Gambar 5. 8. Wordcloud QDA Miner DeepSeek	85
Gambar 5. 9. Bagan Analisis Konten ChatGPT.....	88
Gambar 5. 10. Hubungan Sentimen dan Dominant Factor ChatGPT.....	88
Gambar 5. 11. Bagan Kategori Analisis Konten DeepSeek.....	91
Gambar 5. 12. Hubungan Sentimen dan Dominant Factor DeepSeek	92
Gambar 6. 1. Tren Rating Mingguan ChatGPT	98
Gambar 6. 2. Tren Rating Mingguan DeepSeek.....	98
Gambar 6. 3. Jumlah Ulasan per Label Sentimen ChatGPT	99
Gambar 6. 4. Jumlah Ulasan Mingguan per Label Sentimen DeepSeek	99

DAFTAR TABEL

Tabel 2. 1. Studi Sebelumnya	8
Tabel 2. 2. Studi Sebelumnya	9
Tabel 2. 3. Studi Sebelumnya	10
Tabel 2. 4. Prediction dan Actual Class	13
Tabel 2. 5. Perbandingan Metode	14
Tabel 4. 1. Kolom Scraping Data Ulasan	24
Tabel 4. 2. Kolom Scraping Info Aplikasi	25
Tabel 4. 3. Perbandingan Hasil Terjemahan Bahasa	31
Tabel 4. 4. Hasil Cleaning Data Duplikasi Aplikasi	32
Tabel 4. 5. Hasil Cleaning Tahap 1 Aplikasi	33
Tabel 4. 6. Hasil Cleaning Tahap 2 Aplikasi	34
Tabel 6. 1. Jumlah Pengguna atau Unduh Aplikasi	100