

**ANALISIS SENTIMEN KUALITAS APLIKASI DISCORD
MENGUNAKAN ALGORITMA NAIVE BAYES DAN SUPPORT
VECTOR MACHINE**

SKRIPSI SARJANA SISTEM INFORMASI



Oleh :

Ichsan Dicki Fareza

217006516029

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI DAN INFORMATIKA
UNIVERSITAS NASIONAL
2025**

**ANALISIS SENTIMEN KUALITAS APLIKASI DISCORD
MENGUNAKAN ALGORITMA NAIVE BAYES DAN SUPPORT
VECTOR MACHINE**

SKRIPSI SARJANA

Karya ilmiah sebagai salah satu syarat untuk memperoleh gelar Sarjana Sistem Informasi dari
Fakultas Teknologi Komunikasi dan Informatika



Oleh :

Ichsan Dicki Fareza

217006516029

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI DAN INFORMATIKA
UNIVERSITAS NASIONAL**

2025

HALAMAN PENGESAHAN

TUGAS AKHIR

ANALISIS SENTIMEN KUALITAS APLIKASI DISCORD
MENGUNAKAN ALGORITMA NAÏVE BAYES DAN SUPPORT
VECTOR MACHINE



Ichsan Dicki Fareza

217006516029

Dosen Pembimbing

A handwritten signature in blue ink, consisting of a series of loops and strokes, representing the name of the supervisor.

(Ir. Endah Tri Esti Handayani, MMSI.)

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

ANALISIS SENTIMEN KUALITAS APLIKASI DISCORD MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MACHINE

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber informasi atau acuan yang dicantumkan sebagaimana mestinya.



Jakarta, 28 Februari 2025



Ichsan Dicki Fareza

217006516029

LEMBAR PERSETUJUAN REVIEW AKHIR

Tugas Akhir dengan judul :

ANALISIS SENTIMEN KUALITAS APLIKASI DISCORD MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MACHINE

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diujikan pada Sidang Review Akhir Semester Ganjil 2024-2025 pada tanggal 26 Februari Tahun 2025

Dosen Pembimbing



Ir. Endah Tri Esti Handayani, MMSI.

NIDN. 0314106802

Ketua Program Studi



Dr. Andrianingsih, S.Kom., MMSI.

NIDN. 0303097902

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Ichsan Dicki Fareza
NPM : 217006516029
Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika
Program Studi : Sistem Informasi
Tanggal Sidang : Rabu, 26 Februari 2025

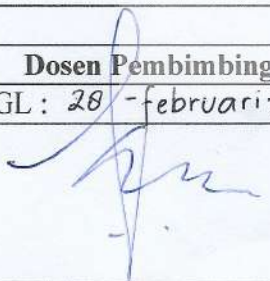
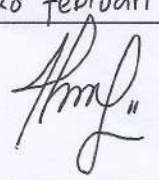
JUDUL DALAM BAHASA INDONESIA :

ANALISIS SENTIMEN KUALITAS APLIKASI DISCORD MENGGUNAKAN
ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MACHINE

JUDUL DALAM BAHASA INGGRIS :

ANALYSIS SENTIMENT QUALITY OF DISCORD APPLICATION USING NAÏVE
BAYES ALGORITHM AND SUPPORT VECTOR MACHINE

TANDA TANGAN DAN TANGGAL

Dosen Pembimbing	Ka. Prodi	Mahasiswa
TGL : 28 -februari-2025	TGL : 28 Februari 2025	TGL : 28 februari 2025
	 Dr. Andrianiyngsih - MMS	

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Allah SWT atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul **“Analisis Sentimen Kualitas Aplikasi Discord Menggunakan Algoritma Naive Bayes dan Support Vector Machine”**. Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan program Sarjana, Jurusan Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika.

Dalam proses penyusunan skripsi ini, penulis mendapatkan banyak bantuan dan dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya terutama kepada dosen pembimbing Tugas Akhir, Ibu Ir. Endah Tri Esti Handayani, MMSI atas waktu, tenaga, dan pikiran dalam membimbing penulis selama penelitian dan penulisan skripsi. Penulis turut menyampaikan rasa terima kasih kepada :

1. Allah SWT yang telah memberikan nikmat panjang umur, kesehatan, kemudahan dan kelancaran dalam menyelesaikan penelitian serta penulisan skripsi ini dengan sebaik mungkin.
2. Kedua orang tua penulis, Bapak Sulardi dan Ibu Kiki Fauzeny yang telah memberikan dukungan dalam berbagai bentuk yang tak terhitung.
3. Ibu Ir. Endah Tri Esti Handayani, selaku Dosen Pembimbing Skripsi.
4. Seluruh dosen pengajar di Program Studi Sistem Informasi FTKI maupun dosen di Program Studi lain yang telah memberikan banyak ilmu.
5. Teman – teman seangkatan yang sudah memberikan bantuan serta dukungan.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun. Semoga skripsi ini dapat memberikan manfaat bagi pembaca serta menjadi kontribusi yang berguna dalam bidang analisis data.

Jakarta, 14 Februari 2025

Ichsan Dicki Fareza

ABSTRAK

Discord merupakan salah satu aplikasi komunikasi berbasis suara, video, dan teks yang banyak digunakan oleh berbagai komunitas. Namun, seiring dengan meningkatnya jumlah pengguna, kualitas layanan aplikasi ini menjadi aspek penting yang perlu dianalisis. Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap ulasan pengguna Discord yang diambil dari Google Play Store, dengan menerapkan algoritma Naïve Bayes dan Support Vector Machine (SVM). Data yang digunakan berjumlah 5000 ulasan berbahasa Indonesia yang dikumpulkan melalui teknik web scraping. Proses analisis sentimen melibatkan tahapan preprocessing data, seperti pembersihan teks, normalisasi, tokenisasi, dan stemming. Label sentimen ditentukan menggunakan TextBlob dalam tiga kategori positif, netral, dan negatif. Hasil evaluasi menunjukkan bahwa algoritma SVM memiliki akurasi yang lebih tinggi dibandingkan Naïve Bayes, dengan nilai akurasi masing-masing sebesar 88% dan 78%. Dengan demikian, algoritma SVM lebih efektif dalam mengklasifikasikan sentimen pengguna terhadap aplikasi Discord. Hasil penelitian ini diharapkan dapat memberikan wawasan bagi pengembang Discord dalam meningkatkan kualitas layanan aplikasi mereka.

Kata kunci : Analisis Sentimen, Naïve Bayes, Support Vector Machine, Discord

ABSTRACT

Discord is a voice, video, and text-based communication application that is widely used by various communities. However, as the number of users increases, the quality of service of this application becomes an important aspect that needs to be analyzed. This research aims to perform sentiment analysis on Discord user reviews taken from Google Play Store, by applying Naïve Bayes and Support Vector Machine (SVM) algorithms. The data used amounted to 5000 Indonesian-language reviews collected through web scraping techniques. The sentiment analysis process involves data preprocessing stages, such as text cleaning, normalization, tokenization, and stemming. Sentiment labels were determined using TextBlob in three categories positive, neutral, and negative. The evaluation results show that the SVM algorithm has higher accuracy than Naïve Bayes, with accuracy values of 88% and 78%, respectively. Thus, the SVM algorithm is more effective in classifying user sentiment towards the Discord application. The results of this study are expected to provide insight for Discord developers in improving the quality of their application services.

Keywords : *Sentiment Analysis, Naïve Bayes, Support Vector Machine, Discord*

DAFTAR ISI

HALAMAN JUDUL	i
KATA PENGANTAR.....	ii
ABSTRAK	iii
ABSTRACT	iv
DAFTAR ISI.....	v
DAFTAR GAMBAR.....	viii
DAFTAR TABEL	x
BAB I.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah.....	3
1.3 Rumusan Masalah.....	4
1.4 Tujuan Penelitian	4
1.5 Batasan Masalah.....	4
1.6 Kontribusi.....	5
BAB II	6
2.1 Studi Literatur	6
2.2 Landasan Teori.....	9
2.2.1 Analisis Sentimen	9
2.2.2 Natural Language Processing.....	9
2.2.3 Machine Learning	10
2.2.4 Web Scraping	10
2.2.5 Naive Bayes	10
2.2.6 Support Vector Machine	11
2.2.7 Python	12
2.2.8 Google Colab	12
2.2.9 Google Play Store	13
2.2.10 Discord	13
2.2.11 Kualitas	13

BAB III	14
3.1 Lokasi Penelitian & Timeline Penelitian	14
3.1.1 Lokasi Penelitian	14
3.1.2 <i>Timeline</i> Penelitian	14
3.2 Fokus Penelitian	14
3.3 Populasi & Sampel Penelitian	14
3.4 Teknik Pengumpulan Data	15
3.4.1 Studi literatur	15
3.4.2 Jenis Data	15
3.4.3 Web Scraping	15
3.5 Desain Penelitian	15
3.6 Metode Evaluasi Hasil	17
3.6.1 Confusion Matrix	17
3.7 Perangkat Penelitian	18
BAB IV	20
4.1 Pengumpulan Data	20
4.2 Preprocessing Data	22
4.2.1 <i>Data Cleaning</i>	22
4.2.2 Normalisasi	23
4.2.3 Stopword Removal	25
4.2.4 Tokenizing	26
4.2.5 Stemming	27
4.3 Data Preparation	29
4.3.1 Translating	29
4.3.2 Labeling	30
4.4 Implementasi Algoritma	32
4.4.1 <i>Splitting</i> Data dan TF - IDF	32
4.4.2 Perhitungan Manual Algoritma	33
4.4.3 Penerapan Colab <i>Naive Bayes Classifier</i>	38
4.4.4 Penerapan Colab <i>Support Vector Machine</i>	40
4.5 Analisis Hasil dan Evaluasi	42
4.5.1 Distribusi Sentimen	42
4.5.2 <i>Wordcloud</i> Modus Kata	43

4.5.3	Analisis Hasil Sentimen Positif.....	44
4.5.4	Analisis Hasil Sentimen Netral	46
4.6	Deployment Model Klasifikasi Sederhana	51
BAB V		55
5.1	Kesimpulan	55
5.2	Saran.....	55
DAFTAR PUSTAKA		56



DAFTAR GAMBAR

Gambar 3.1 Desain Penelitian.....	16
Gambar 4.1 Package name Discord	20
Gambar 4.2 Proses Scrapping Data.....	21
Gambar 4.3 Hasil Web Scrapping.....	21
Gambar 4.4 Cek Duplikat	22
Gambar 4.5 Cek Nilai Null	22
Gambar 4.6 Data Cleaning.....	22
Gambar 4.7 Hasil Data Cleaning	23
Gambar 4.8 Normalisasi	24
Gambar 4.9 Hasil Normalisasi.....	24
Gambar 4.10 Stopword Removal.....	25
Gambar 4.11 Hasil Stopword Removal	26
Gambar 4.12 Tokenizing	26
Gambar 4.13 Hasil Tokenizing	27
Gambar 4.14 Stemming	28
Gambar 4.15 Hasil Stemming	28
Gambar 4.16 Translating	29
Gambar 4.17 Hasil Training.....	30
Gambar 4.18 Labeling	31
Gambar 4.19 Hasil Labeling	31
Gambar 4.20 Splitting Data & TfidfVectorizer	32
Gambar 4.21 Hasil Pembagian Data Latih & Uji	33
Gambar 4.22 Model Naïve Bayes Multinomial	38
Gambar 4.23 Classification Report Naive Bayes.....	39
Gambar 4.24 Model Support Vector Machine	40
Gambar 4.25 Classification Report Support Vector Machine	41
Gambar 4.26 Distribusi Sentimen.....	42
Gambar 4.27 Wordcloud Modus Kata	43
Gambar 4.28 Wordcloud Modus Kata Sentimen Positif.....	44

Gambar 4.29 Bar Chart Modus Kata Sentimen Positif	44
Gambar 4.30 Network Chart Hubungan Kata Sentimen Positif	45
Gambar 4.31 Wordcloud Modus Kata Sentimen Netral	46
Gambar 4.32 Bar Chart Modus Kata Sentimen Netral	46
Gambar 4.32 Network Chart Hubungan Kata Sentimen Netral.....	47
Gambar 4.34 Wordcloud Modus Kata Sentimen Negatif	48
Gambar 4.35 Bar Chart Modus Kata Sentimen Negatif	48
Gambar 4.36 Network Chart Hubungan Kata Sentimen Negatif.....	49
Gambar 4.37 Confussion Matrix Naïve Bayes	50
Gambar 4.38 Confussion Matrix SVM	51
Gambar 4.39 Hasil Prediksi Sentimen Model NB & SVM	52



DAFTAR TABEL

Tabel 2.1 Studi Literatur	6
Tabel 3.1 Timeline Penelitian	14
Tabel 3.2 Confussion Matrix	18
Tabel 3.3 Perangkat Penelitian	19
Tabel 4.1 Sampel Data Latih NB	33
Tabel 4.2 Sampel Data Uji NB	33
Tabel 4.3 Ulasan Positif	34
Tabel 4.4 Kata Relevan Pada Ulasan Positif	34
Tabel 4.5 Ulasan Negatif	34
Tabel 4.6 Kata Relevan Pada Ulasan Negatif	34
Tabel 4.7 Frekuensi Kata Dalam Ulasan Positif	35
Tabel 4.8 Frekuensi Kata Dalam Ulasan Negatif	35
Tabel 4.9 Sampel Data SVM	36
Tabel 4.10 Bag of Words	37
Tabel 4.11 Contoh Vektor Fitur	37
Tabel 4.12 Contoh Vektor Biner Ulasan Baru	38
Tabel 4.13 Hasil Uji Prediksi Kalimat Baru	52